T

T

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES B—BIOLOGICAL SCIENCES

VOL 137

1950

T

T

CONTENTS

No. B 889. 28 November 1950

Address of the President
Sir Robert Robinson, at the
Anniversary Meeting, 30 November 1949

Awards of Medals, 1949

The COPLEY MEDAL is awarded to Professor GEORGE CHARLES DE HEVESY, For.Mem.R.S., for his distinguished work on the chemistry of radioactive elements and especially for his use of isotopes as tracers in the study of biochemical problems.

Hevesy was one of the last to join the distinguished company of discoverers of elements in the classical tradition. In 1923, in collaboration with Coster, he established the occurrence of the element with atomic number 72 in zirconia minerals, and called it hafnium. This was shown to be a close analogue and constant comparison of zirconium, but a method of separation by chemical means was devised. The atomic weight was found to be 178·6, and the X-ray and optical spectra were fully described.

Here he wrote the last chapter of a volume and now we turn to the beginning of a new book.

When working under Rutherford in Manchester, Hevesy turned his failure to separate radium D from lead to good advantage. He recognized that the identity in chemical properties of radium D, and other radioactive isotopes of lead, with ordinary lead, made it possible to use these radioactive isotopes as indicators to follow the behaviour of lead in chemical processes, and in great detail owing to the extreme sensitivity of the methods of detection of radioactivity.

The first application of the idea was made with Paneth in 1913 at the Radium Institute of Vienna, and it was followed by a large number of interesting researches of a physico-chemical nature.

The use of radioactive isotopes as indicators or 'tracers' in biological processes was initiated in 1923 by Hevesy's studies on the uptake and distribution of lead in bean plants, using radioactive lead, radium D or thorium B, as indicators. After placing bean plants in solutions of ordinary lead nitrate containing small amounts of radium-D nitrate, the distribution of the lead was followed by determining the radioactivity of the ash from the different parts of the plants. This work may truly be said to have marked the opening of a new chapter in biochemistry. Furthermore, it established a pattern for the numerous subsequent researches by himself and by others which were to follow the discovery of artificial radioactive elements and the development of methods for the separation of certain stable isotopes.

Hevesy was one of the first to appreciate the potential biochemical importance of Urey's discovery of deuterium in 1932.

In the following year, by experiments with fish placed in water containing added D_2O, he and Hofer showed that there is a rapid exchange between environmental water and that in the body, and also that there is an exchange between the hydrogen

of the environmental water and labile hydrogen atoms in the tissue constituents. In 1934, by experiments on human subjects in which the density of the urinary water was compared with that of water ingested, they established the very important fact that at the low concentrations of D_2O present in ordinary water the body does not discriminate between D_2O and H_2O. It was found in the course of this work that the average time during which a water molecule remains in the body is 13 ± 1.5 days.

In 1937 Hevesy and his co-workers again broke fresh ground by their use of the radioactive isotope of phosphorus, ^{32}P, as a tracer in studies of the metabolism of phosphorus compounds. For example, ^{32}P-labelled phosphate was administered to rabbits, and its uptake by various tissues and its rate of excretion were determined. It was found that the average time during which a phosphorus atom remains in the body is thirty days. Having established by *in vitro* experiments that there is no direct exchange between organic ester phosphate and inorganic phosphate, Hevesy and his co-workers were able to determine the 'turn-over' rates for certain organic phosphorus compounds in the body by isolating these compounds at intervals after the administration of labelled phosphate and estimating their content of radio-active ^{32}P.

Work on the uptake of labelled phosphate by bone did much to establish that the whole of the skeleton is in a state of dynamic equilibrium with the constituents of the body fluids. By studies on lactating animals Hevesy and his collaborators were able to show that ^{32}P is incorporated into the caseinogen of milk within a few hours. They obtained ^{32}P-labelled adenosine triphosphate enzymatically from labelled inorganic phosphate, and clarified the role of this compound as a phosphate donor in carbohydrate metabolism. Further, their studies on the phosphatides in the liver and blood plasma supported the view that the plasma phosphatides are synthesized in the liver.

The spark of the Manchester days has kindled a flame that can illumine all that is dynamic in biochemistry.

A ROYAL MEDAL is awarded to Sir GEORGE PAGET THOMSON for his distinguished contributions to atomic physics and especially for his investigation of the wave properties of the electron.

During a period of more than thirty years, Thomson has made many notable contributions to experimental and theoretical physics over a wide range of subjects.

In 1914–18 he developed important aspects of the science of aeronautics and took part in experimental flights in order to investigate problems of interest to the air services. His results were communicated in Government reports, and at the end of the war he wrote a book *Applied Aeronautics*.

Early work in the laboratory was, with filial piety, connected in some way or another with positive rays, and in this field he disclosed a number of important results.

But Thomson's most distinguished contribution to knowledge is his demonstration of the wave nature of the electron. Shortly after, but independently of Davisson's discovery of the diffraction of electrons by single crystals, Thomson found that

a narrow pencil of swift electrons, after transmission through a thin film of poly-crystalline matter, produced, on a photographic plate or fluorescent screen, a pattern of rings analogous to optical halos or to the Debye-Scherrer rings well known in the corresponding experiment with X-rays. By simple and beautiful experiments he was able to show that the patterns could only be explained on de Broglie's revolutionary theory of the wave nature of matter, and he proved, within the experimental accuracy of 1%, that the wave-length associated with the electron must be h/mv, as required by the theory.

In addition to bringing conclusive proof of this remarkable duality in the behaviour of matter, these experiments opened out a new and fruitful field of research which has important practical applications. Thomson showed that electron diffraction provided an eminently suitable method for the study of the structure of surfaces. As a result it is now possible to investigate how the structure of the surfaces of metals is changed by mechanical, thermal or chemical treatment, and thus to extend by a powerful method the information about the structure of matter which can be obtained by the complementary method of X-ray diffraction.

In later years Thomson's interests have extended to nuclear physics and important contributions on some aspects of this subject, including the effects of cosmic rays and the properties of mesons, have come from his laboratory.

Thomson's work has been in the direct line of progress; it has given us a new conception and a new tool for research. In both respects it has notably widened our horizons.

A ROYAL MEDAL is awarded to Professor RUDOLPH ALBERT PETERS for his distinguished researches in biochemistry and in particular his discovery of the role of vitamin B_1 in tissue metabolism.

The dramatic recovery of an animal suffering from some particular vitamin deficiency when given a minute amount of that vitamin, coupled with the observation that the curative dose must be repeated at intervals to prevent a relapse to avitaminosis, had drawn attention to the possibility that vitamins were either catalysts or played some important role in essential enzyme system.

But it was the work of Peters on the brain tissue of vitamin B_1-deficient pigeons, begun in 1929, which showed for the first time that the latter suggestion was indeed correct. In 1931 he found that polyneuritic symptoms in such pigeons could be correlated with a general deficit in lactic acid metabolism and that this was most marked in the lower regions of the brain. Comparing normal and avitaminous brain tissue he concluded that B_1 must play some essential part in tissue respiration, namely in the glucose-lactic acid system. Following the elucidation by R. R. Williams and others of the chemical structure of B_1 (aneurin) he was able to show that the vitamin, in its phosphorylated form as co-carboxylase, was part of the pyruvic oxidase system and was thus an essential factor in carbohydrate metabolism.

In 1937 Peters found that pyruvic oxidase, on which he was working in connexion with vitamin B_1, was poisoned by small amounts of arsenite; later, with Sinclair and Thompson, he showed that it was sensitive also to certain organic arsenicals. This

led him to suggest that the toxicity of such substances was due to their action on the functional thiol groups of this enzyme.

It transpired that the more potent vesicants, containing the group $AsCl_2$, combined with two thiol groups of kerateine (Stocken & Thompson), and this suggested the use of a synthetic dithiol such that an arsenic atom, two sulphur atoms, and two or three carbon atoms might be locked together in a stable ring system. The object would be to protect the enzyme by exhibiting, in effective mass, an alternative anchor for trivalent arsenic. A new dithiol, now known as B.A.L., was selected as the most promising representative of its type, and this was found to shield the pyruvic-oxidase system from arsenicals and even to be capable of reversing a toxic action when introduced at a later stage. B.A.L. has been found to be effective against arsenical vesicants in man and to show a beneficial action against arsenical dermatitis in a substantial number of cases. It has recently been found to be useful also in acute mercurial poisoning and in the dermatitis that may follow administration of compounds of gold.

There can be little doubt that the discovery of B.A.L. by Peters and his associates represents an advance of considerable importance in practical therapeutics. It was clearly no chance observation but the result of a logical development conforming to the valid principles of scientific method.

We can make for Peters the proud claim that he was the first to elucidate the role of a vitamin in animal metabolism.

The DAVY MEDAL is awarded to Professor ALEXANDER ROBERTUS TODD for his analytic and synthetic studies in organic chemistry and biochemistry with special reference to vitamins and nucleosides.

In early collaborative work on the bile acids with Borsche the Wieland-Windaus constitutions current at that time were shown to be inadequate. Later, at Oxford, Todd made a very substantial contribution to the synthesis of diglycosidic anthocyanins, and he also helped to determine the structure and to effect the syntheses of mould products of the anthraquinone series, such as helminthosporin and cynodontin. He also established the main lines of the constitutions of flavoglaucin and auroglaucin.

His work on aneurin was of great value; he took a part in the determination of the structure of vitamin B_1 and developed syntheses, one of which is now employed for the manufacture of the vitamin. In the course of this investigation he established the structure of thiochrome and proved it by synthesis. He determined, in parallel with Fernholz & Karrer, the structure of vitamin E and effected its synthesis.

His studies of cannabinol, the active principle of hashish, were also crowned by synthesis.

From this stage on, his work has been characterized by increasing and individually characteristic originality. He has provided the last details of the structure of all the natural purine and pyrimidine nucleosides, and has synthesized them. A resounding success was the synthesis of A.D.P. and A.T.P., or adenosine triphosphate, and the artificial preparation of cozymase has been brought near to its final stages. This work involved much attention to pyrimidine chemistry, to some aspects of the

carbohydrates, to the processes of phosphorylation, and to the properties of phosphoric esters and anhydrides. It is of quite outstanding merit, a milestone in the progress of biochemistry.

Other problems which Todd has illuminated, and often solved, concern the nature of the specific germinating factor for *Striga hermonthica* (this turned out to be D-xyloketose); the nature of the factor produced by solanaceous plants and which is required for the hatching of eelworms; the mechanism of the hardening process of insect cuticle. He has recently cleared up the chemistry of a series of *bis*-isoquinoline alkaloids and started an intensive study of insect pigments, especially the remarkable colouring matters of Aphididae, which exhibit such interesting colour changes.

Todd has proved himself a master of the strategy and tactics of research, and the great school which he directs at Cambridge has enhanced the prestige of British chemistry throughout the world. His achievements have already added greatly to our knowledge of organic chemistry and biochemistry, and still more they have indicated sure pathways for further advance towards the solution of some of the central problems of biology. Can it be an exaggeration to describe in this way work which is essential for a better understanding of at least one aspect of the fundamental chemistry of nucleo-proteins?

The SYLVESTER MEDAL is awarded to Professor LOUIS JOEL MORDELL for his distinguished researches in pure mathematics, especially for discoveries in the theory of numbers.

Although he is also a powerful analyst, Mordell has always been primarily an arithmetician. Indeed, he was for long almost the only British mathematician of whom this could be said, and, if this is no longer true, it is mainly the result of his own teaching and example.

His most enduring interest has been in the theory of indeterminate equations. This is a subject which has attracted many great mathematicians, but it is one of supreme difficulty and new ideas emerge only at rare intervals. One of Mordell's greatest achievements was his proof in 1921 of the 'finite basis' theorem for the rational solutions of a cubic equation $f(x, y) = 0$ in two variables. This asserts, roughly speaking, that all the rational solutions can be derived, by a systematic process, from a finite number of them. The theorem represented a great advance in the theory of indeterminate equations, and has inspired a great deal of work by other mathematicians all over the world. In other papers, from his earliest to his most recent, Mordell has developed a variety of methods which throw new light on many different types of indeterminate equations.

Another field in which Mordell has made great advances is that of the geometry of numbers. One general problem here is that of finding the best possible inequality for the minimum of an algebraic form of a given type, when the variables take integral values. The results for a binary quadratic form are classical. The binary cubic form was considered by several mathematicians of the last century, but the final result was found by Mordell as late as 1940. The geometrical interpretation of the problem involves a particular non-convex region in the plane, and the methods which Mordell devised for its solution proved to be applicable to other non-convex

regions. They enabled problems to be solved which would formerly have been considered unapproachable.

A third region in which Mordell has shown his mastery is the theory of the elliptic modular functions and their applications to the theory of numbers. In particular, he used the theory to give a new treatment of the representation of numbers as sums of squares, and to prove the validity of some of Ramanujan's conjectures.

Mordell's output relating to these topics and to numerous individual problems is characterized by the subtlety and fertility of the methods he has discovered. He has had a profound influence on the development of the theory of numbers, and has been widely recognized as one of the most eminent mathematicians of our time, both for the importance of his own researches and for his inspiration of the work of others.

The HUGHES MEDAL is awarded to Professor CECIL FRANK POWELL for his distinguished work in experimental physics and especially for the discovery of mesons and their transformations.

Powell's early researches were concerned with the properties of ions. These led to the investigations concerning fundamental particles and atomic nuclei using the photographic plate technique, for which he and his school of experimental physics at Bristol have a world-wide reputation.

He was instrumental in bringing about a marked improvement in the quality of the sensitive material especially developed for research in nuclear physics, and using these improved plates, has discovered a new fundamental particle, the π meson, of which the mass is 280 times that of the electron. The negative π mesons are captured by atomic nuclei, causing the nuclei to disintegrate. The positive π mesons decay spontaneously, when at rest, into μ mesons of 205 electron masses, whose existence had been established by other means, and these in turn into a positive electron and presumably a neutrino.

These discoveries relating to mesons have proved to be of profound importance in unravelling the complicated phenomena of cosmic radiation in the earth's atmosphere and have thrown new light on the theory of nuclear forces.

High-energy nuclear disintegration can so far only be studied by means of cosmic rays, and Powell has also been responsible for several remarkable discoveries in this field.

All these and other nuclear transformations have been recorded in a most striking and beautiful way in the photographs of Powell and his school.

In him we salute a master of advanced technique which has made possible discoveries that are certainly of transcendent natural philosophical importance. They may well prove to be also of great practical significance in unsuspected directions in relation to the physical system as a whole.

I wish to thank Members of Council for the great help they have given me in the preparation of notes on the Medallists.

I do not propose on this occasion to expatiate on the matters raised in the Report of Council or to allude to any of the more general topics, which concern our interest.

The reason is that several important developments are in process, and a statement made at the present stage would have little value.

As the result of the ballot has been announced I am entitled to hope that the position in regard to the most important of current activities, namely, accommodation of the scientific societies, will be clearer in a year's time, and that I shall be able to announce definite progress.

There is continuous, though slow, improvement in the speed of publication of papers in the Society's *Proceedings* and *Transactions*. An important factor is the time taken by referees, and the majority of Fellows have responded with corresponding willing action to the request that referee forms should be returned within three weeks of the receipt of a paper.

The Secretaries and Publication Committee hope for still further acceleration of this stage of the work.

On behalf of the Fellows of the Society I take this opportunity of expressing to the British Thomson-Houston Company our warm thanks for their generous action in financing the installation of fluorescent lighting in our rooms.

THE SIGNIFICANCE OF THE STERINOIDS

There are certain sections of organic chemistry of such a specialized nature that only a few persons have a real mastery of their intricacies. Examples are the chemistry and technology of azo-dyes, or of anthraquinone derivatives and related polycyclics. And certainly the chemistry, biochemistry and physiology of the sterinoids cannot be comprehended fully except by those who devote undivided attention to these subjects. It would clearly be impossible to describe any part of this vast field at all adequately in the few minutes at my disposal, and I propose to do no more than glance at a few of the headings of developments in the last twenty years. My object is to justify a plea for an even greater effort in research and for international co-operation such as was achieved during the war in the penicillin field.

Although cholesterol is a constituent of all animal cells there was no inkling of its physiological importance when Windaus began his investigations in 1903, and this was also true in 1912, the approximate date of commencement of Wieland's work on the related bile acids.

The molecular structures were studied as interesting organic chemical problems, curiosities, attractive because of their difficulty and the unique character of the group. Nevertheless, the impressive accumulation of facts in this period established many of the fundamental relations and laid the foundation for further progress.

The formulae that were regarded as correct for several years were:
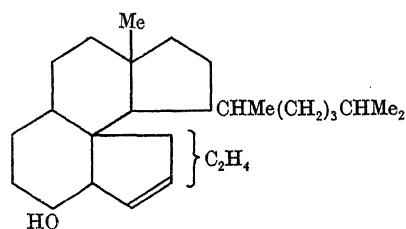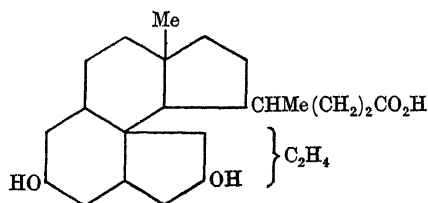
cholesterol (1928)   desoxycholic acid (1928)

Various attempts to accommodate the group C_2H_4 in these structures were unsuccessful, so that it may be said that no full and satisfactory constitution was advanced. It was gradually realized that the problem was quite unsolved in regard to the precise nature of the polycyclic portion of the molecule, though the evidence for the side-chains, including the relation of cholic acid to cholesterol, was conclusive. As mentioned below the connexion between ergosterol and vitamin D was appreciated in 1926, and the increase of interest in this aspect of the subject probably led to the next step in advance.

In 1932 Bernal examined ergosterol crystals by means of X-rays and pointed out that the molecular dimensions could not be reconciled with the Wieland-Windaus structures. He found that the molecule was longer and thinner than those theories suggested, and this recalled similar conclusions of Adam & Rosenheim, drawn in 1929 from the study of surface films. In some cases these authors assumed a tilt in the molecules, but they were in general disposed to state that their results could not be explained on the basis of the accepted formulae.

The nettle was grasped by Rosenheim & King who recalled the work of Diels (1927) on the dehydrogenation of certain sterol derivatives, whereby the hydrocarbon chrysene was formed along with other products. They decided to see what would happen if the formation of chrysene (I) was regarded as significant and not, as heretofore, the result of deep-seated decomposition. This led them in May 1932 to advance the formula (II) for desoxycholic acid.
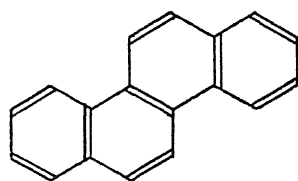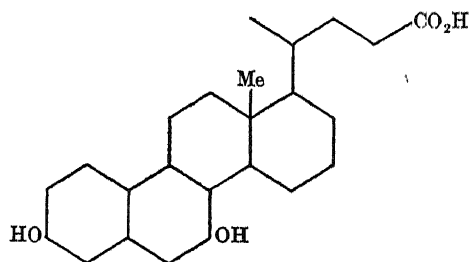
(I)   (II)

A structure of this type, applied to ergosterol was found by Bernal to harmonize with his view of the overall dimensions of the molecules.

A few months later in the same year, Wieland & Dane, and also Rosenheim & King, made the final modification to (III) for desoxycholic acid; the corresponding cholesterol structure is (IV) [cholic acid is (III) with a third hydroxyl in position 7].
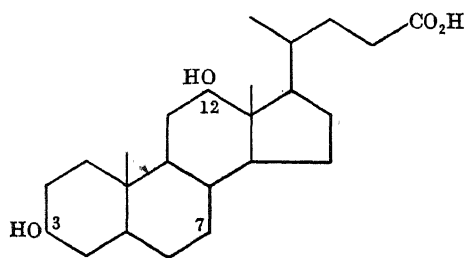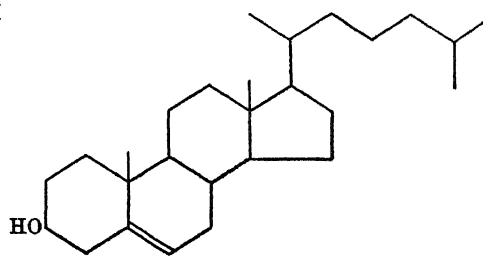
(III)   (IV)

Then followed a busy period in which the assumptions made were tested and found in every case to be valid.

Wieland had shown how to degrade the long side-chain step by step, and Tschesche applied a similar process to an acid (V) obtained by opening ring *A*.

(V)   (VI)
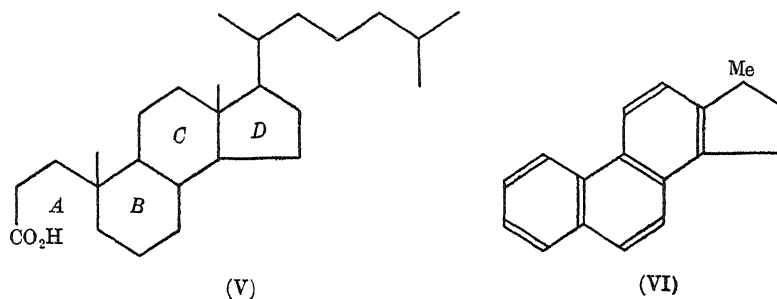
His results followed the scheme $R.CH_2.CH_2.CO_2H$ (V) $\rightarrow R.CH_2.CO_2H \rightarrow R.CO_2H$

and the properties of this last acid showed it to be of the form $\begin{matrix} C \\ \diagdown \\ C \\ \diagup \\ C \end{matrix} C - CO_2H$ and thus

disclosed the position of the side-methyl or angle-methyl group.

In addition to chrysene, Diels obtained a hydrocarbon $C_{18}H_{16}$, for example by the dehydrogenation of cholesteryl chloride. This was considered to be (VI), the methyl having wandered into the cyclopentane ring from the angle-position (Cook & Hewitt). This formulation was proved to be correct by synthesis due to Kon, Harper & F. C. J. Ruzicka. An earlier and doubtless successful synthesis by Bergmann & Hillemann was unfortunately not crowned by conclusive comparison with the Diels hydrocarbon. Several other hydrocarbons obtained by dehydrogenation processes have been investigated and synthesized, but by far the most interesting is that known as methylcholanthrene (VIII) because its formation confirms the position of the side-chain and of one of the hydroxyls of cholic acid as well as of the main skeleton.

Wieland & Dane (1933) degraded desoxycholic acid to 12-ketocholanic acid (VII) and thence through dehydro*nor*cholene to (VIII).

The last process was independently effected by Cook & Haslewood whose interest was already aroused in 1932 when Kennaway & Cook discussed the possibility of obtaining, from the bile acids, carcinogens analogous to others which they had discovered in the benzanthracene series. Cook & Haslewood demonstrated the constitution of methylcholanthrene by degradation to a tetracyclic compound which they synthesized, and Fieser & Seligman (1935–6) synthesized methylcholanthrene itself in unambiguous fashion.

In passing it may be remarked that there is no positive evidence to connect methylcholanthrene with the etiology of cancer; it has been looked for, and not found, as a constituent of normal and of cancerous tissues.·

Thus the long-known facts and newly acquired knowledge fell neatly into place in the light of the new structures and further confirmatory evidence was derived from the total synthesis of oestrogens, namely, equilenin by Bachmann, Cole &

Wilds (1939) and oestrone by Anner & Miescher (1948). The close relation of these substances to cholesterol was always thought to be highly probable but not proven until 1940 when Inhoffen forged the last links in the chain of transformations connecting the two series.

(VII)

(VIII)
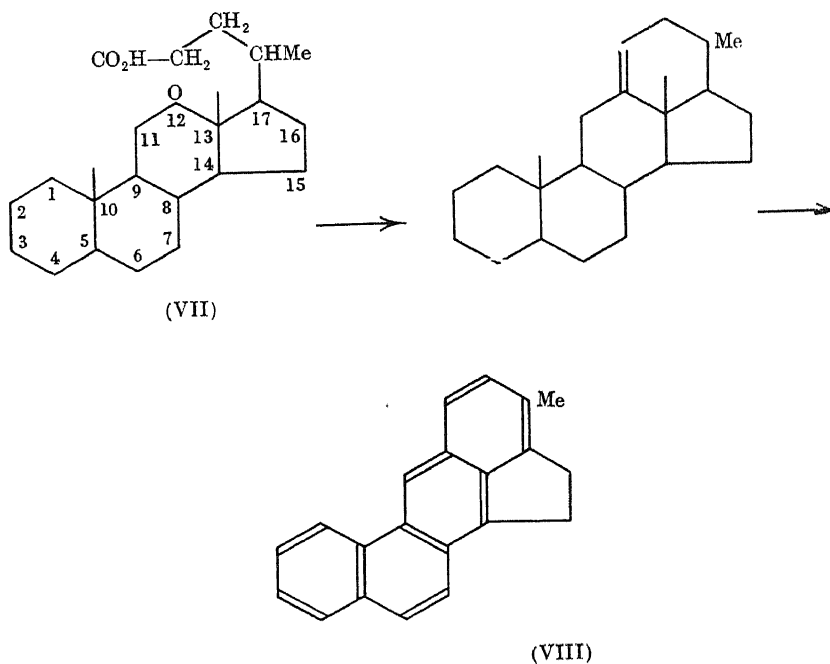
The synthesis of cholesterol itself might well be thought a hopeless quest when it is considered that the formula contains seven asymmetric carbon atoms implying 2^7 stereoisomerides. But one of the doublings is accounted for by resolution into enantiomorphs, and at the worst the problem resolves into six bifurcations in the synthetic route, at each of which some progress must be possible along the right path. We have set ourselves this task and have adopted two methods symbolized as $ABC \rightarrow D$ and $A \rightarrow BCD$, indicating the ring to be added last. Both routes have been developed but we have gone further along the former path. Following it, J. W. Cornforth has provided the first synthetical proof of the correctness of the carbon skeleton in rings A, B and C.

The tricyclic diketone (IX) was obtained by Reich (1945) as a degradation product of desoxycholic acid by way of several intermediates. The Reich diketone can also be obtained from a keto-acetate (X) which was isolated by Köster & Logemann from the debrominated products of the oxidation of cholesterol acetate dibromide.

Naphthalene has been transformed into (IX) through nineteen stages and the diketone was obtained in optically active form stereochemically identical with Reich's product.

Although (IX) is obtainable from (X) the reverse operation has not yet been effected. Nevertheless, (X) is much the more readily accessible as a relay, and we are therefore studying the addition of ring D to it. The first possible bifurcation has given two substances to H. Holtermann and one of them must possess the correct

orientation. This leaves only one more fork to be negotiated so as to reach the full
tetracyclic system in correct stereoisomeric form. We shall be unlucky if the right
road is barred here or if it proves impossible to convert (IX) into (X). The long
ascent is all but finished at the tetracyclic stage. This constitutes a veritable plat-
form, we can walk round to it, and tackle the few remaining pitches whenever we
feel disposed to do so. We are confident that the total synthesis of cholesterol can
be realized along these lines.

(IX)   (X)
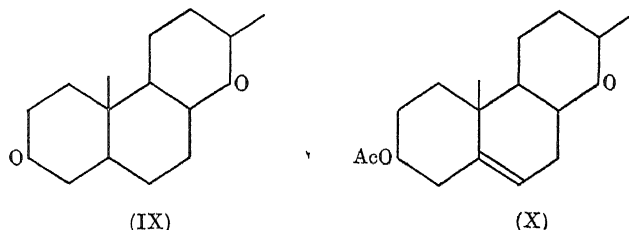
A host of other sterinoids have fallen, or are gradually falling, into line. The
constitution of ergosterol, for example, has been demonstrated by the masterly
researches of Windaus and of Heilbron. It is impossible to do more than mention
the existence of the phytosterols, such as sitosterol and stigmasterol, the very
complex saponins, sapogenins, and heart poisons, the sterinoid alkaloids, and so on.
All these groups present highly interesting and difficult problems and already have
a most extensive literature.

Many of the plant poisons are glycosides and even the sugars derived from them
are unique. One of the simplest, obtained from the purple foxglove, is digitoxin,
which yields on hydrolysis, digitoxigenin (XI) and three molecules of digitose (XII)
which has a *d*-ribose type of configuration (*d*-desoxy-altrose).

$$CH_3(CH.OH)_3CH_2.CHO$$

(XI)   (XII)
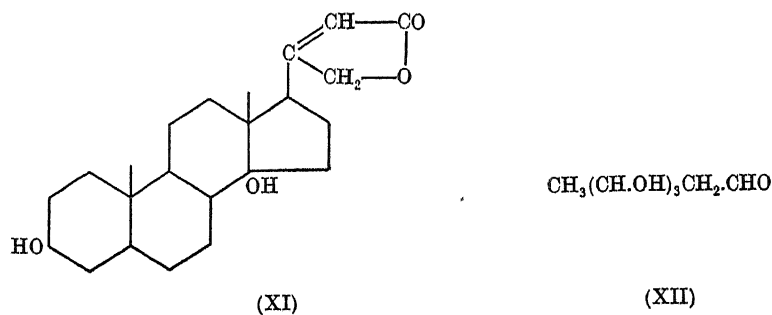
Another peculiarity of some of these sugars is that they include *O*-methyl groups
of true ether function.

Throughout the chemical work on sterols and their derivatives the problem of
stereochemical configuration has been prominent. Much had been done by con-
sideration of transformations, relations and analogies and this was greatly sup-
plemented and extended by X-ray study of crystals. The most exact work in this
field has been carried out by Mrs Hodgkin (née Crowfoot) and her collaborators.
Their complete study of cholesteryl iodide (Carlisle & Crowfoot 1945) confirmed
conclusions drawn from earlier work and clarified all the stereochemical relations of

the molecule. In cholesterol the hydroxyl group and the two angle-methyls lie on the same side of the general plane of the structure. That there is no Walden inversion in the conversion of cholesterol to its iodide may be inferred from the crystal analyses of the iodo-nitrobenzoates of calciferol and lumisterol.

To-day, we have all but a full knowledge of sterol configuration and the results are correlated with rotatory powers as well as with physiological activity.

Vitamin D

Following up the knowledge that rickets could be benefited by sunlight or by addition of certain fish oils to the diet, the active constituent in the oils was concentrated and recognized as a vitamin. Hess, and also Steenbock, found in 1924 that exposure of inactive oils to sunlight or to ultra-violet radiation enhanced their antirachitic properties. Furthermore the pro-vitamin was located in the sterol fraction of the oils. It was found that cholesterol was not this pro-vitamin but that irradiation of ergosterol under suitable conditions produced an antirachitic substance which was termed vitamin D. Intensive study of the transformation products of ergosterol by groups of workers at the National Institute for Medical Research and at Göttingen led to the isolation of the pure vitamin, called calciferol. Then as a result of structural investigations by Heilbron, and by Windaus, and their respective collaborators, the remarkable transformation of ergosterol (XIII) into calciferol (XIV) was established.

C_9H_{17} C_9H_{17}

CH_2

HO   HO

(XIII)   (XIV)
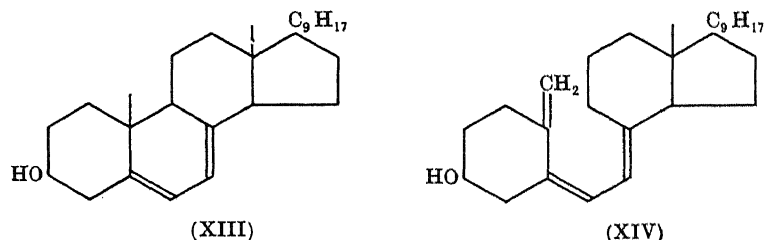
In addition, detailed information about intermediate stages and subsequent changes was garnered and it must be emphasized that this development was possible only as a superstructure on the firm foundations of the constitutional studies of cholesterol and ergosterol. Apart from the side-chain the characteristic feature of the latter is the extra double bond (Δ^7) in ring B and Windaus prepared a cholesterol analogue, namely, 7-dehydrocholesterol and found that it also gave an antirachitic substance on irradiation. This was later recognized as the true naturally occurring vitamin, now designated as D_3. It has been isolated from tunny liver oil by Brockmann and identified, as the 3:5-dinitrobenzoate, with the product from irradiated 7-dehydrocholesterol ((XIII) with side-chain C_8H_{17}). The method of preparation of the pro-vitamin has been greatly improved and it has been found in Nature, for example in the skin of pigs.

Oestrogenic, androgenic, and progestational hormones

No attempt to cover this field even by way of summary can be made.

It will suffice to notice certain features that connect with the general sterol narrative. The isolation of the first oestrogen, namely, oestrone (XV) by Doisy

(August) and Butenandt (October) occurred in 1929 and it will therefore be appreciated that the constitutional study was largely independent of that of the sterols. The problem was solved by the work of Marrian, of Butenandt, and of Cook from about 1932–4 and the outcome was one of give-and-take with sterol chemistry. The discovery of oestrone was quickly followed by that of other oestrogens; oestriol by Marrian in 1930, equilenin (XVI) by Girard in 1932. Reference has already been made to the synthesis of (XV) and (XVI).
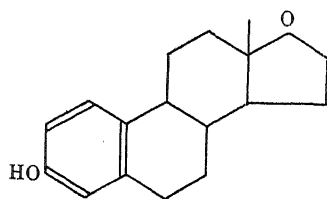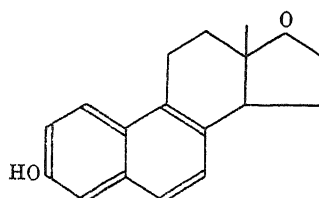
(XV)   (XVI)

These hormones were at first obtained from the urine of pregnant women and later from that of pregnant mares. In 1934 Zondek found that the urine of stallions was a still richer source, but this surprising result is peculiar to equines. These hormones have also been found in plant material; oestrone in palm kernels and oestriol in female willow flowers. The subterranean clover, so important to Australian agriculture, contains oestrogens which have apparently not yet been identified. The oestrogenic property is, however, not very specific and is exhibited by a considerable range of synthetic compounds of which stilboestrol (XVII), hexoestrol, its dihydro-derivative, and dieneoestrol, a dehydro-derivative, are the best known (Doods and collaborators).
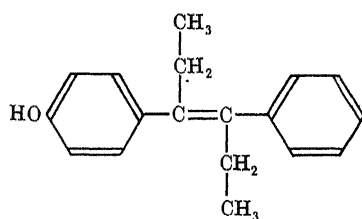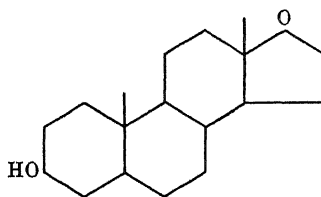
(XVII)   (XVIII)

The natural and also the synthetic oestrogens find many applications in therapeutics and both types are manufactured on a considerable scale.

Androsterone (XVIII) was isolated by Butenandt in 1931 from urine of normal males and he almost guessed the correct constitution in 1932. Two years later Ruzicka entered the field with the highly important discovery that he could oxidize away the side-chain of sterol and bile acid derivatives and leave a carbonyl group in its place: $:CH.C_8H_{17} \rightarrow :CO$. The oxidation of cholestanyl acetate gave a weak androgen, similar to, but not identical with, androsterone. Use of a stereo-isomeride, epi-cholestanyl acetate afforded androsterone itself in very small yield.

In 1935 Laqueur isolated a very active hormone, namely testosterone (XIX) from testes of steers and this has become the standard androgen in medical practice. It

is manufactured by a development of Ruzicka's method in which the starting point is the dibromide of cholesteryl acetate, which is oxidized by chromic acid. This was reported in 1935 from four different laboratories.

(XIX)
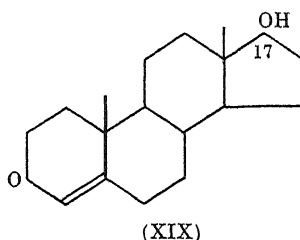
The final stage is the reduction of a carbonyl group in position 17 which is best carried out by a method due to Mamoli; treatment with a sugar-yeast fermenting solution.

Until quite recently no simple synthetic androgenic substance was known, although Birch & Mukherji have described the preparations of XX from hexoestrol by an ingenious method and this substance is under examination for its biological properties.

(XX)   (XXI)
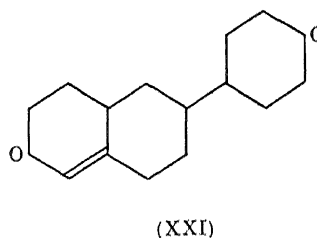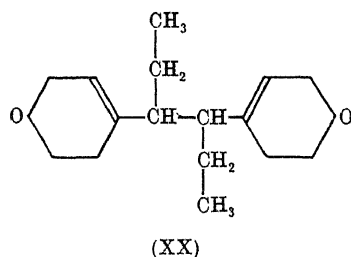
Wilds, Shunk & Hoffman (September 1949) have prepared the diketone (XXI) which is stated to possess approximately 1/200th of the androgenic activity of testosterone in the comb test in day-old chicks. This discovery is of great interest because it gives new encouragement to the search for synthetic analogues of hormones of therapeutic importance.

Progesterone (XXII) is a hormone of the corpus luteum which controls pregnancy. Following indications by W. M. Allen (1932) it was independently isolated in pure form in four laboratories in 1934. The best partial syntheses are from stigmasterol, from various sapogenins, or from a by-product of the oxidation of cholesteryl acetate dibromide.

Hormones of the adrenal cortex

Following the indications (1929) that extracts of the gland could prolong the life of adrenalectomized animals, Kendall, Wintersteiner & Pfiffner, and Reichstein have (since 1935) described the isolation of some twenty-eight crystalline substances from the adrenal cortex. This magnificent effort cannot be summarized but the constitutions of all the more important of these substances have been demonstrated by transformations or partial syntheses. Corticosterone (XXIII) was the first active substance to be isolated, first by Reichstein and a little later by Kendall. Like some

of its congeners it is concerned with regulation of carbohydrate metabolism and electrolyte balance but to discuss these relations in a few words would be impossible. Indeed it is not surprising that some doubt still exists in regard to the precise function of individual hormones in this complex group. The adreno-cortical hormones have been used in the treatment of Addison's disease and surgical or wound shock.

CO.CH₃ structure (XXII), CO.CH₂OH structure (XXIII), CO.CH₂OH structure (XXIV)
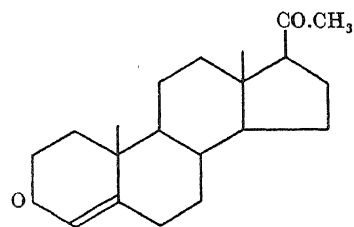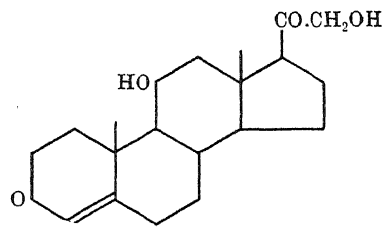
(XXII)   (XXIII)
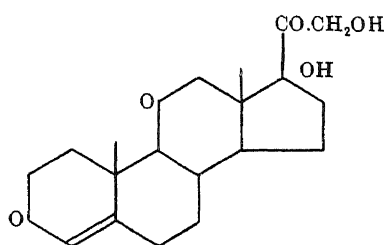
(XXIV)

At the moment the centre of the stage is occupied by Kendall's substance E (Reichstein's substance F) (XXIV) which is now called cortisone.

Five years ago Selye concluded from his experiments that the adrenal cortex may play an important role in the pathogenesis of rheumatic and rheumatoid conditions in man, and Bassi & Bassi (1946) obtained beneficial effects in chronic rheumatoid arthritis by the administration of adrenal-cortical extracts. Cortisone has now been found by Hench, Kendall, Slocumb & Polley (1949) to alleviate the symptoms of rheumatoid arthritis in a quite dramatic fashion. However, the effect is only temporary and further administration of the hormone is necessary after a certain period. We cannot yet know the length of time during which the improvement can be sustained and undesirable side reactions have appeared in some cases. There are said to be six million rheumatoid arthritics in the United States alone and the provision of cortisone or some effective substitute is therefore a matter of great urgency. But very large quantities will be required and it is doubtful whether the amount needed could conceivably be made by partial synthesis from sterinoid material of natural origin.

At the present time the starting point for successful syntheses is desoxycholic acid which is carried through some thirty-five stages in accordance with brilliant work of Reichstein, and of Sarett, to which many others have contributed by way of modification and improvement of yield. By a remarkable *tour de force* the chemists

of Merck and Co., Rahway, N.J., U.S.A., have made some hundreds of grams of cortisone, but colossal effort was involved. It is claimed that the number of stages would be halved if sarmentogenin could be used as the starting point; this aglycone of a cardiac glycoside from a species of *Strophanthus*, already bears an oxygen atom in position 11. But this material is not available in quantity and, even if it were, the partial synthesis is still too long for the purpose in view. The more promising lines would appear to be (a) total synthesis, (b) synthesis of an effective analogue, or (c) synthesis of a substitute.

The feasibility of finding analogues cannot be assessed in advance. Undoubtedly this idea will be pursued in more than one quarter. Total synthesis is attractive in this case but demands new ideas and new methods; there is a great field for investigation.

The substitute may be found along lines indicated by L. H. Li (First Congress of Biochemistry, Cambridge, 1949). It is first necessary to realize that the hormones already mentioned are released on the arrival of messengers from a higher centre. Thus the anterior pituitary discharges gonadotropic hormones into the blood. These stimulate the gonads to the production of hormones of sterinoid nature. The gonadotropic hormone, studied in this country by Rimington, among others, has protein character. Similarly the anterior pituitary produces adrenocorticotropic hormone (A.C.T.H.) which stimulates the adrenal cortex and which Hench *et al.* found equivalent to cortisone in the treatment of rheumatoid arthritis. The essence of Li's discovery is that a partial hydrolysate of this protein contains a peptide with a small number of amino-acids (variously stated as 6 to 12) and that this substance is active in the same sense as A.C.T.H. Clearly the synthesis of such a substance might be practicable on a really large scale.

Work on the mode of action of cortisone and A.C.T.H. is of the greatest interest. A recent paper of Selye draws attention to the fact that adrenal cortical hormones are of two main types, mineralo-corticoid and gluco-corticoid and that cortisone and A.C.T.H. are in the latter category. The idea of a balance between these types may explain some apparent anomalies such as the fact that arthritic symptoms do not necessarily appear at a low level of activity of the adrenals.

Finally, mention must be made of some very interesting work by Rittenberg & Bloch on the biogenesis of cholesterol. The sterol is synthesized by surviving rat liver slices *in vitro* from labelled acetate and subsequently degraded in various ways. The results indicate that all the carbon atoms can be derived from acetate. It will soon be possible to attribute to each of the carbon atoms its origin from a methyl or carboxyl of acetic acid.

In spite of, or better, because of, all these brilliant achievements it cannot be supposed that more than a small part of the wonder of the sterols and their derivatives has thus far been brought to light.

Reverting to the tropic hormones of the anterior pituitary, is should be noted that these are by no means the only examples. There seems to be a fairly general and adaptable mechanism whereby the appearance of a protein stimulates the formation

or release of various steroid hormones at different sites in the body. There must be some common factor in these processes and it may not be so far from our reach as we imagine.

There is a widespread belief that steroid hormones are intimately connected in some way with malignancy and for this reason alone research in the field must be vigorously pursued. All the topics we have touched upon reach out into the unknown and all are of profound significance for the progress of biological science and hence for the health and happiness of mankind.

There is a good case for the international organization of a supreme effort which could conveniently be directed in the first instance towards following up the cortisone-A.C.T.H. clues.

Those who control UNESCO have wisely decided to initiate certain International Laboratories, but it is doubtful whether any of the subjects so far advanced, important as they undoubtedly are, should be preferred to a concerted investigation of the chemistry, biochemistry and endocrinology of the sterinoids.

This is not the place or the time to consider the details, which can be adjusted to suit the different circumstances of the participants, but it is already clear that centralization in terms of bricks and mortar would prove difficult.

An organization, with ample funds to be used for the promotion of research in existing centres, and capable of effecting full liaison between them, is the first objective.

The organization and work of the Division of Biochemistry and General Nutrition of C.S.I.R.

By H. R. MARSTON,* *Chief of Division*

[Plates 1 to 7]

INTRODUCTION—ORGANIZATION AND ADMINISTRATION

In allotting me the task of relating to you something of the organization and work of the Division of Biochemistry and General Nutrition of the Council for Scientific and Industrial Research, I trust it is not your pleasure to have me dwell on the first part of this title—for a description in measured terms of a substance so tenuous is beyond me. Organization (in its official sense) is not readily discernible to those of us who are gathered together in this laboratory. It is there, but it is never obtrusive.

We are an organ of a body which, although nurtured by Government funds, is free from the vegetative nervous system of the public service administration. We are situated within the precincts of a university, and enjoy the amenities and intellectual discipline of academic life without having to shoulder too great a burden of teaching. Our reference is to conduct research (with a quiet aside that applied science may remain healthy and productive only when thoroughly fertilized with more fundamental studies). Here is a grand experiment in Government administration.

The burden of responsibility entailed is apparent. This, together with a simple faith that practical application will flow freely and naturally from more complete understanding of the underlying phenomena, directs our efforts; and a common curiosity welds us, without loss of individuality, into a team. Our accounts are kept in a way which is acceptable to the Treasury.

The team is composed of individuals whose initial training has been either in physiology or in chemistry—eight senior research men, about twice this number of younger men (the heart of the team), and about four times as many technical assistants (enough to provide all the extra hands and eyes the team may use effectively without being overwhelmed by problems of management). The backbone of the team consists of individuals who have had post-graduate training in British universities and who are imbued with the ideals of which this place is the Mecca.

THE SCOPE OF THE WORK

The meagre fund of exact knowledge of ruminant physiology—which, in light of the importance of sheep and cattle to the well-being of mankind, is rather a depressing heritage—prompted our attention to this general field. Our interest became

fixed on the nutritional biochemistry of the sheep—a convenient and tractable ruminant—and more especially on factors which influence wool production, for here was a virgin field that might appropriately be cultivated in Australia.

There is no great mystery about the nutritional factors which govern wool production by normal, healthy sheep. The quota of the amino-acids absorbed from the intestine that ultimately serves as a substrate for wool production must escape alike the demands of competing syntheses, and of deamination and utilization as a source of energy—modifying influences which vary with the nutritional state of the animal.

A series of experiments in which there was a precise knowledge of the overall energy transactions, of the nitrogen balance, of the amino-acids in the fodder and in the main synthetic products—wool and flesh, and of the rate of wool production, allowed the influence of various states of nutrition on the efficiency with which amino-acids are utilized for wool production to be assessed with satisfactory accuracy. Appraisal of the nutritional limitations imposed by grazing conditions then became feasible and we were able to proceed intelligently with their correction.

The estimation of the overall efficiency of wool production was greatly simplified by a study of the amino-acid constitution of the vegetable proteins which normally provide the substrate from which wool is elaborated, for this revealed that the protoplasmic proteins, apparently of all orders of higher plants, are of extraordinarily similar composition. The biological value of these proteins for wool production is restricted by their potential cystine content to approximately 35 % under ideal conditions. The theoretical nutritional state in which there would be a complete conversion of the limiting essential amino-acid, however, is never closely approached —even under nutritional conditions in which the energy requirements are met with substances other than protein, and there is a minimum of depletion of the substrate of amino acids by competing syntheses, the efficiency with which the highly evolved Merino sheep converts to wool fleece the amino-acids absorbed from its intestine exceeds 10 % only in exceptional instances. Obviously wool keratin is an expensive biological product when compared with the proteins of meat, milk, eggs, etc.

Usually the concentration of protein in pastures falls very short of that necessary to provide for maximum fleece production, and for this reason the grazing Merino rarely expresses its full hereditary propensity—fine wool, more often than not, is mainly the result of suboptimum nutritional conditions.

The experimental work which is proceeding on the thermodynamics of food utilization, designed primarily to provide a sound basis for maintaining flocks over periods of drought, must be passed over with mere mention. Although energy metabolism was the central theme of physiology at the close of last century, it had been abandoned in a state of no little confusion. About ten years ago we built two calorimeters for the study of energy transactions in the sheep and these tools have since been used continuously to provide an essential basis for our nutritional investigations. Discussion of the work on the intermediary metabolism of the simple fatty acids which, as products of microbial dissimilation of carbohydrates, constitute the main and apparently wasteful source of energy for the ruminant, must be set aside, and so must the studies of the nutrition and activity of the symbiotic microflora of

the paunch, upon which the ruminant largely depends. These projects lack neither scientific interest nor basic nutritional importance, but time is limited and I should like to relate in some detail our experiments with trace elements as these have led in a sort of Pilgrim's Progress (which we like to imagine was a logical sequence) to the solution of agricultural problems of considerable economic significance.

TRACE ELEMENT STUDIES

The geochemistry of the very deficient post-Pliocene aeolian soils

A great part of the soils along the seaboard of southern Australia are of aeolian origin—their genesis probably dates from a period of extreme aridity in post-Pliocene times.

The coastal fringes are composed essentially of marine shell fragments, wind-borne into gently rolling dunes that cover unconformably and at considerable depth the underlying strata. These merge imperceptibly through consolidated limestone ridges, intersected by brown soils—residua derived apparently from leaching of the older dunes—into a hinterland where light siliceous sands imposed upon solonized subsoils prevail over relatively huge areas.

Indigenous plant associations on these soils

Most of this terrain is in a zone of secure and adequate rainfall. The calcareous littoral supports a very simple plant association in which two calciphile grasses, *Bromus madritensis* and *Lagurus ovatus*, overwhelmingly predominate (plate 1 a). The siliceous sands constitute the so-called mallee deserts where a low and stunted heath of dwarfed eucalypts, *Banksia* sp., *Xanthorrhea* sp., etc., has become inured to the penury of the deficient environment (plate 1 b).

Desultory attempts to farm these tracts failed, and all but the marginal areas were abandoned as useless for agriculture.

Coast disease and related maladies

The earliest settlers who ventured to graze their flocks on the natural pastures of the littoral found their enterprise hazardous. Sheep confined to these tracts invariably became anaemic and would waste and die, usually within a year.

This was the first problem that attracted us—it provided the clue which led to the solution of the problem of these refractory lands.

A short preliminary study revealed a syndrome very similar to that of a malady which had been observed to affect sheep depastured on specific areas in most countries, and which had been dubbed with many names. Human experience with this malady had been a long one. Its aetiology was quite obscure.

We proceeded on the assumption that the malady had a nutritional origin. Was it then an uncomplicated deficiency of an essential nutrient element, or was it a deficiency imposed by excessive ingestion of calcium? Experiments soon dispelled the latter idea—and incidentally placed in proper perspective the '*Erdalkali*

Alkalizität' theory which, at that time, tended to obscure our conception of the phosphorus metabolism of ruminants.

The former hypothesis was encouraged when we observed that sheep in an advanced stage of the malady would respond dramatically to dosing with a mixture of soluble salts of the heavy metals (Ni, Co, Mn, Zn, Fe, Cu) which apparently are always present as traces in living matter. As the therapeutic effects were strikingly obvious, cobalt was identified as the active constituent of this mixture within a few months, and a year later we had demonstrated unequivocally that 1 mg. Co/day would prevent the onset of the symptoms which invariably affect sheep depastured on the littoral (figures 1, 2 and plate 2).

The inference that those maladies which are characterized by a similar syndrome might also originate from cobalt deficiency proved correct. A year or so after our first announcement, the New Zealand bush sickness which affects sheep on soils derived from acidic volcanic ejecumenta, the Western Australian enzootic marasmus which affects sheep on certain red loams, the vinkish and daising of the Cheviots, the pine of the raised beaches of the Inner Hebrides, the salt sickness of the coral sands of Florida, and similar maladies elsewhere were all demonstrated, by others, to respond dramatically to cobalt.

But the disabilities suffered by sheep depastured on the deficient littoral were of more complex origin. Another syndrome, previously masked by the more rapidly fatal effects of cobalt deficiency became evident in the animals treated with cobalt. This proved to be the result of copper deficiency.

A further four years of experimental observations conclusively demonstrated that adequate supplements of cobalt and of copper would maintain sheep in normal health and productivity while depastured on this terrain; and collateral chemical evidence left no doubt that the malady was the result of a dietary deficiency of both cobalt and copper. Here, then, where the dual deficiency prevails, appropriate treatment would reveal the uncomplicated syndrome of either deficiency (figure 3).

A unique opportunity was thus provided for intensive study of these deficiencies.

Cobalt deficiency

We found it possible to produce the deficiency syndrome under laboratory conditions by confining sheep to concrete pens, and feeding them on rations consisting of hay produced on the deficient terrain, supplemented with washed gluten and cod-liver oil—a diet which, when further supplemented with cobalt and copper, fulfilled their nutritional requirements. This procedure rendered possible a close study of the malady (figure 4).

A sheep exhibits characteristic symptoms when cobalt deficient (plate 2). Its normal alert demeanour changes to one of dull and rheumy-eyed listlessness; its buccal and conjunctival mucosae blanch; its skin pales from the normal ruddy pink to a dull greenish hue, becomes less flexible and in the advanced stages of the malady sufficiently fragile to break on parting the wool. A steady decline in body weight reflects a failing appetite which progresses until the animal dies of inanition. Autopsy reveals little that is distinguishable from hunger oedema—there was little evidence which might provide a clue to the physiological mechanisms involved.
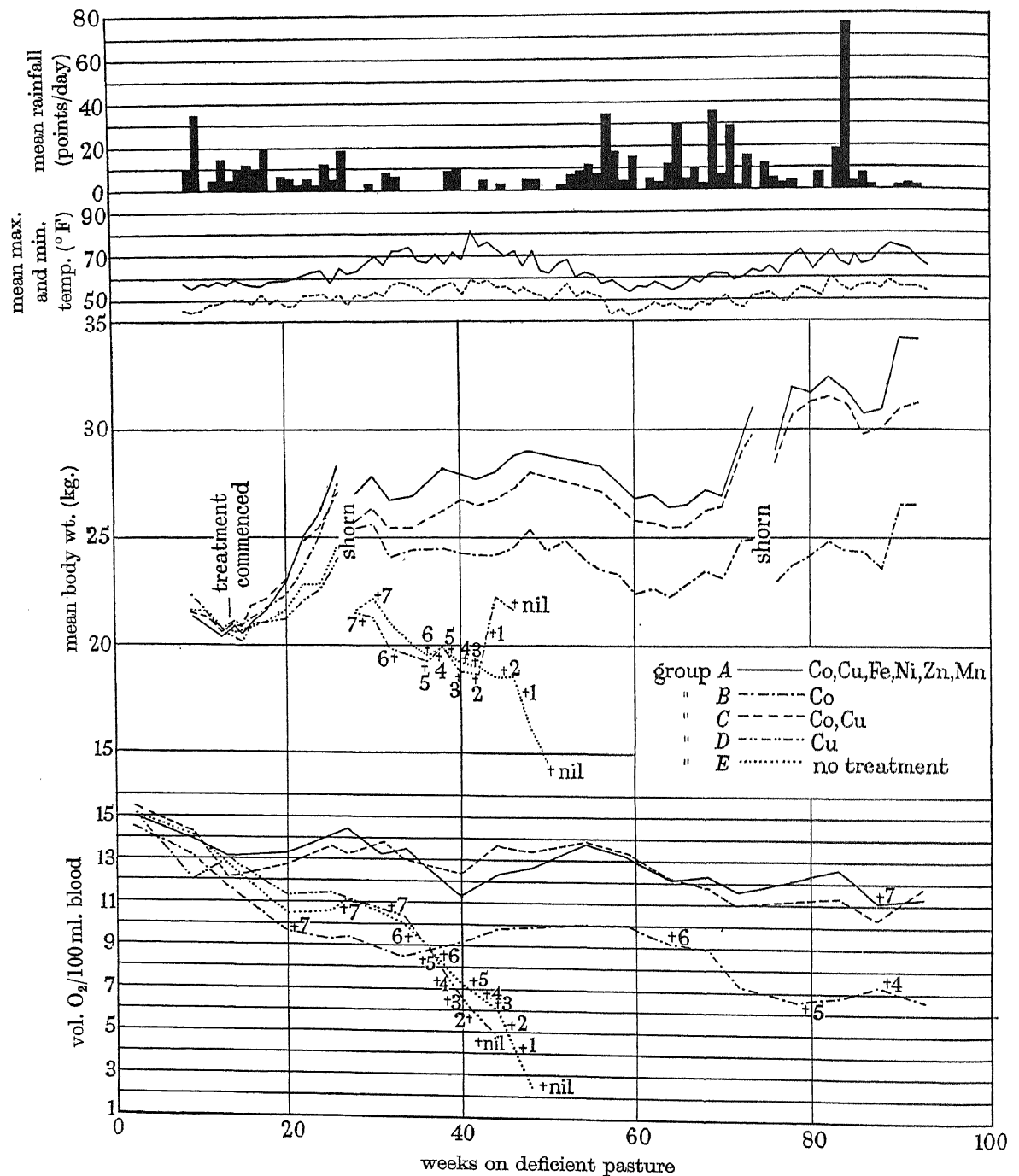
FIGURE 1. The data are from an experiment in which an even flock of ewes (*aet.* 9 months at the beginning of the experiment) was divided into 5 groups each of 8 animals and treated, while depastured on the deficient littoral, by dosing thrice weekly with the elements indicated. The effects of additional cobalt are obvious. Those without it all succumbed within a year. The effects of copper deficiency previously masked by the more rapid effects of cobalt deficiency become apparent in group *B*. Other elements superimposed on cobalt and copper exerted no significant benefit. (From experiments of Marston, Lee & McDonald 1936.)
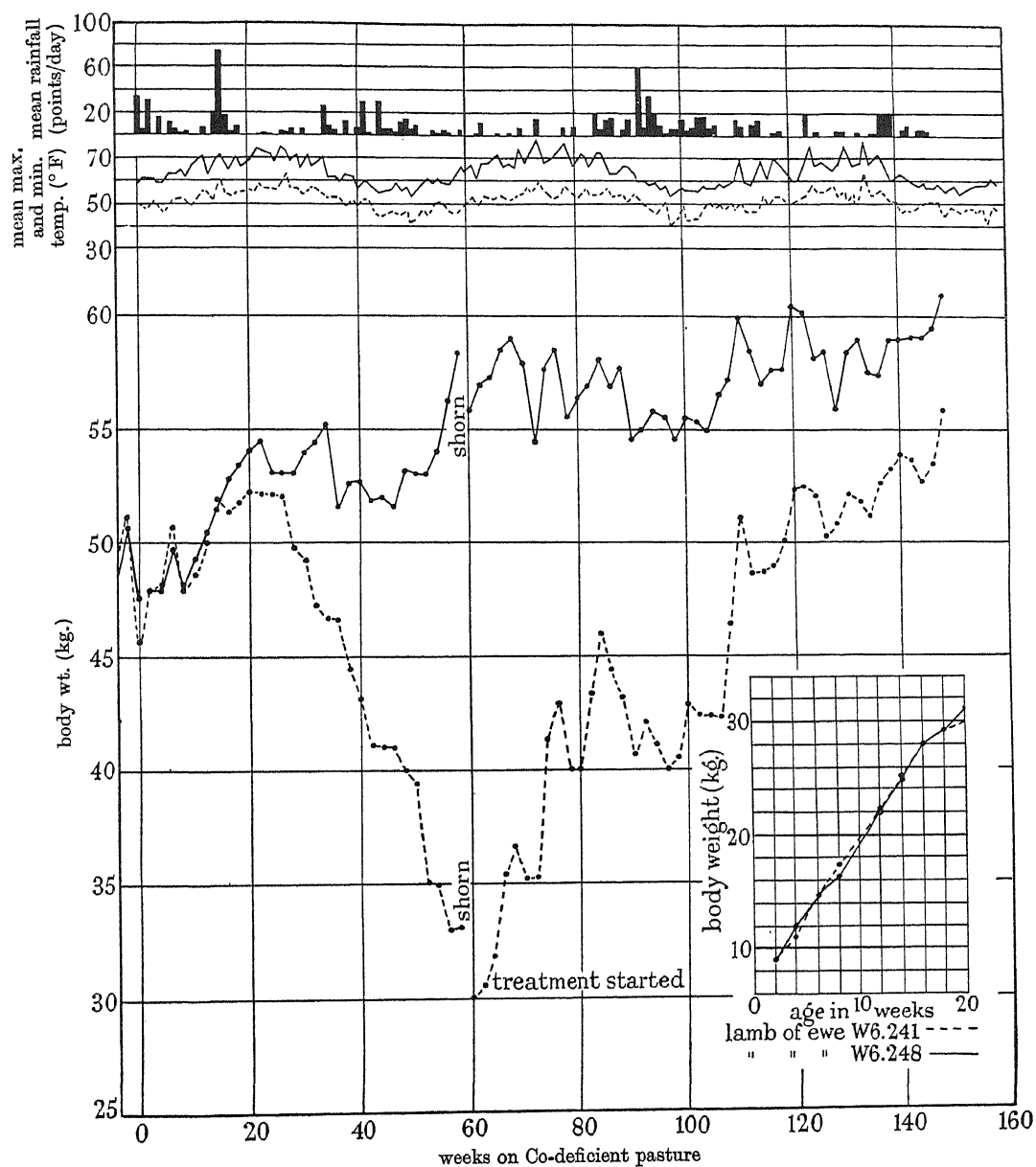
FIGURE 2. The data refer to one of a number of pairs of mature twin Merino ewes which were transferred to the deficient littoral in 1936. Ewe W 6.248 was treated with cobalt throughout the experimental period of 3 years. Ewe W 6.241 was untreated during the first 55 weeks after which a supplement of cobalt sulphate equivalent to 1 mg. Co/day was administered. Both ewes received 10 mg. Cu/day. The supplement of cobalt prevented loss of weight and the appearance of deficiency symptoms in ewe W 6.248. The effect of cobalt administered to ewe W 6.241 while in the terminal stage of the deficiency syndrome (see plate 13) illustrates the typical therapeutic action. There are no striking permanent effects after recovery from an extreme state of cobalt deficiency. The growth rates of the lambs (inset) produced subsequent to mating these ewes at 110 weeks are practically identical. Both received 1 mg. Co/day. ●——● ewe W 6.248 (treated), ●- - -● ewe W 6.241 (untreated).

The anaemia which is a feature of the malady presents a blood picture suggestive of an aplastic condition of the bone marrow (plate 2 b). In the early stages, a reduction in the number of cells rather than changes in their mean size or in their mean haemoglobin content contributes to the low oxygen-carrying capacity. Later, as the state of anaemia worsens, the picture becomes one of a profound large-cell anaemia, somewhat reminiscent of, but quite distinct from that of Addison's pernicious anaemia of man.

All attempts to produce cobalt deficiency in small laboratory animals (rats, mice, rabbits, etc.) have failed—most of the experiments reported in the literature are
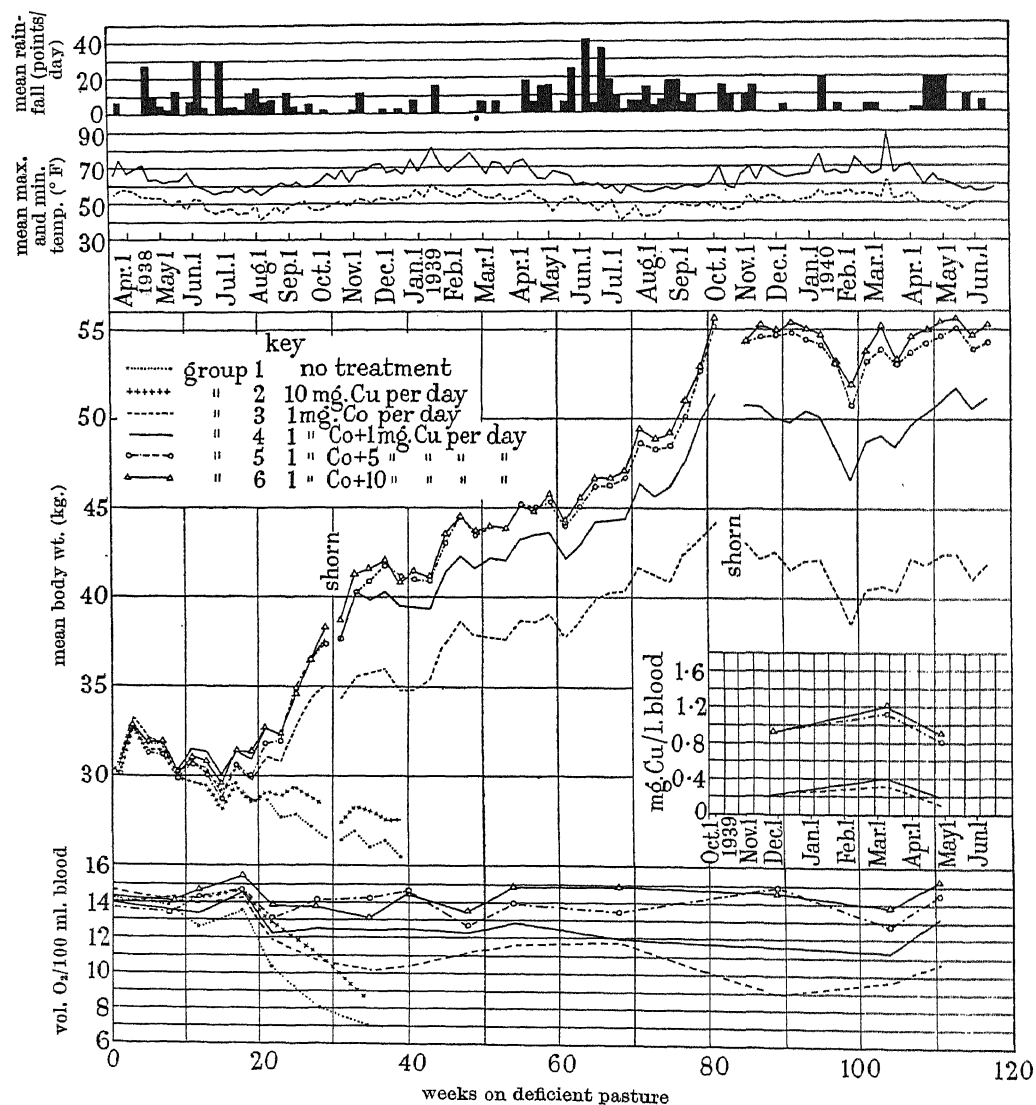
FIGURE 3. The data are from an experiment designed to estimate the quantity of copper necessary to fulfil the requirements of sheep depastured on the deficient terrain. This flock was divided into six groups each of 10 animals and treated as indicated. (From experiments of Marston, Lee & McDonald 1938.)
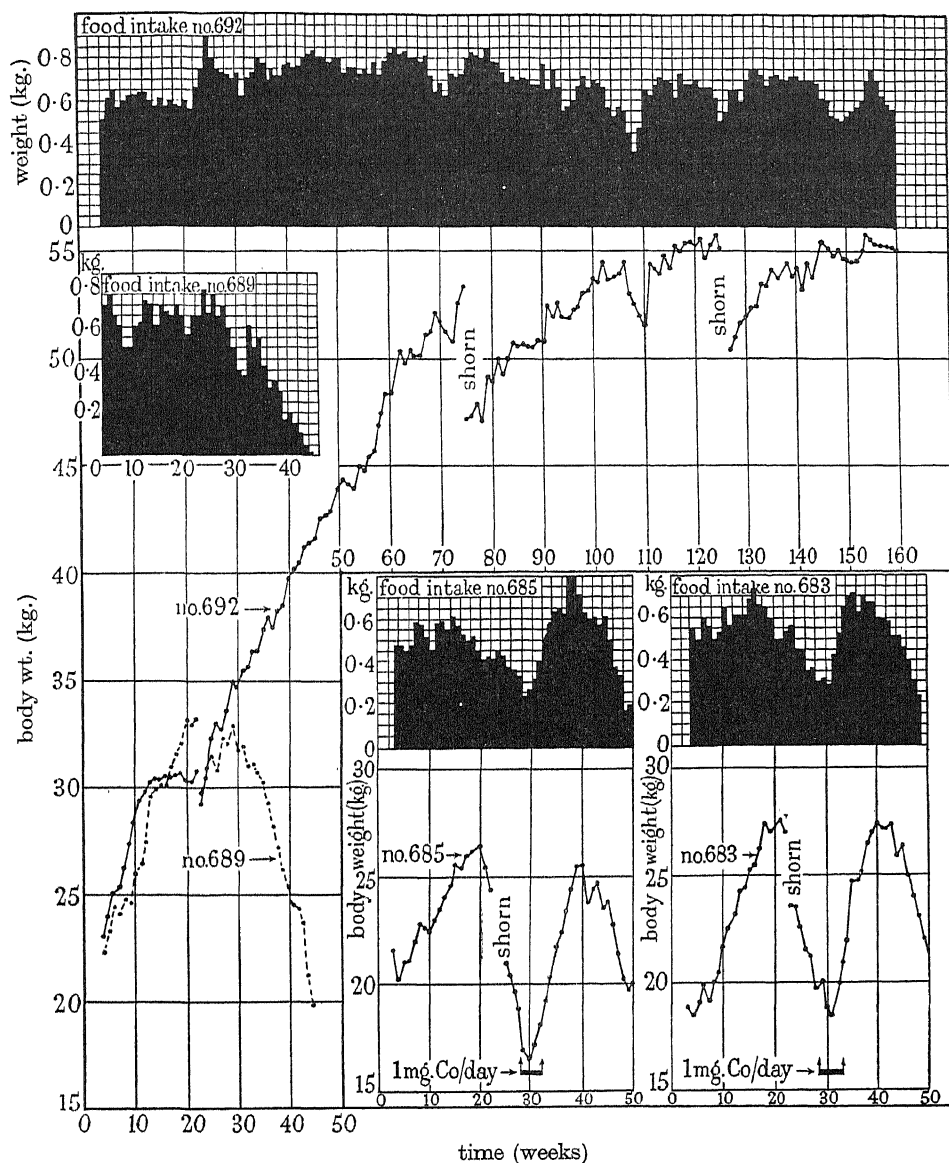
FIGURE 4. Rye hay grown on the deficient littoral supplies adequate nourishment for sheep if provided *ad lib.* and supplemented each day with 100 g. washed gluten, 4 ml. cod-liver oil and 1 mg. Co and 10 mg. Co. Without additional cobalt this diet is inadequate—sheep fed on it develop progressive symptoms of cobalt deficiency which terminate fatally. The data illustrate the effect of the cobalt supplement on sheep that have been pen-fed on this diet. The growth rate of no. 692 which received 1 mg. Co/day from the beginning of the experiment is normal. No. 689, to which no extra cobalt was supplied, developed well-defined symptoms of cobalt deficiency, lost weight and died 44 weeks after being confined to this diet, the course of the syndrome being identical with that observed in the field.

The insets show the effects which supervened when 1 mg. Co/day was provided for 4 weeks to sheep which had developed the deficiency syndrome under these conditions. Note effects on food intake which is represented as mean daily intake of food/week. The rye hay contained 0·02 μg. Co/g. dry wt.; the gluten 0·02 μg. Co/g. dry wt. The water provided for the animals was distilled. (From unpublished experiments of Marston & Peirce 1938–9.)

unconvincing, however, as they have been of short duration and the chemical claims have been of doubtful significance; but we have now the third filial generation of rats that have been sustained, with elaborate precautions and meticulous chemical control, on a diet containing < 0·03 parts of cobalt per million dry weight, upon which they remain normal and healthy. Sheep would certainly succumb to the deficiency within a year if confined to fodder containing this concentration of cobalt.

The amount of cobalt necessary to overcome the fatal consequences of the deficiency is certainly very small. The equivalent of 1 mg. Co/day which was arbitrarily adopted in the first series of experiments proved on further study to be in large excess of the actual needs of sheep depastured on the deficient littoral where the intake in the fodder varies between 20 and 30 μg. Co/day—the equivalent of 0·1 mg. Co/day was sufficient to prevent entirely the appearance of any untoward symptoms, *provided that it was administered frequently*. In these experiments the positive controls were dosed thrice weekly; correspondingly larger doses given at intervals of two weeks failed to prevent the onset of deficiency symptoms, and equivalent doses administered at intervals of 6 weeks merely delayed their fatal termination.

The fact that sheep and cattle are apparently the only animals which suffer the consequences of cobalt deficiency while grazing on the deficient terrain—marsupials abound, the rabbit population if unchecked assumes plague proportions, and horses do exceptionally well there—suggests that the ruminant is either extraordinarily wasteful with cobalt or is unique in its greater demand for this element. Our observations support the latter hypothesis.

In our first experiments with sheep transported from normal pastures to the deficient littoral we had observed the deficiency syndrome develop to its fatal conclusion in some whose livers, when analyzed subsequent to death, had retained as much cobalt as is found in the livers of animals grazed elsewhere on normal pastures. Later, we established depots in the sheep by introducing relatively massive quantities of slightly soluble cobalt salts under the skin, and observed no benefit from this procedure. On investigating the reason for this, we obtained unequivocal evidence that if cobalt is to be effective it must be ingested—when injected into the blood stream it is valueless to the animal (figure 5) although the cobalt concentration in the organs may by this means be increased at least ten-fold above that found in normal healthy sheep. Clearly then, cobalt exerts its action primarily either in the lumen of the alimentary canal or during its passage through the wall. The site of this activity appears to be at a level above the duodenum, for study of the fate of Co^{60} injected intravenously revealed that a considerable proportion of cobalt introduced in this way is voided in the faeces; it is not secreted in the saliva but finds its way via the bile into the intestine from which it is not reabsorbed.

An important function of cobalt in the ruminant might thus concern the symbiotic flora of the paunch upon which the nutrition of the ruminant depends. If the demands of these micro-organisms are large compared with the requirements of the animal itself, then the strikingly dissimilar behaviour between ruminants and other herbivora might be explained.

Attention has now been fixed on the essential nature of cobalt for the nutrition of animals other than ruminants, by the announcement a few months ago of the isolation of Minot's anti-pernicious anaemia factor from liver, and of the discovery, both in the United Kingdom and in the United States, that this bright red, crystalline material is an association complex of cobalt.

These findings already suggest that the quantity of cobalt required for the normal physiological function of the bone marrow is minute, as $< 10 \mu g$. of this factor (equivalent to $< 0.4 \mu g$. Co) introduced parenterally has been reported to correct for at least 7 days the dyscrasia of human pernicious anaemia.

Deficient pastures, which are never completely devoid of cobalt (they contain between 0.02 and $0.03 \mu g$. Co/g. dry wt.), may provide to all grazing animals sufficient cobalt for the elaboration of this factor—possibly by interaction with Castle's extrinsic factor in the acid gastric contents—but leave unfulfilled, in the ruminant, the requirements of other physiological mechanisms (those of the symbiotic micro-flora for example) which demand greater quantities of cobalt for their normal activities.

The profound anaemia that supervenes on cobalt deficiency in the ruminant has certain features in common with other macrocytic anaemias of nutritional origin. But its origin is apparently not due primarily to a deficiency of the A.P.A. factor. Our current experiments suggest that it does not respond either to massive amounts of liver extracts administered parenterally or to similar therapy with the pure A.P.A. factor. Nor will it respond to dosing *per os* with pteroyl-glutamic acid. In so far as our experience extends, cobalt administered *per os* (as a soluble ionizable salt) is the only effective agent which will prevent the onset of the deficiency syndrome in sheep depastured on the deficient tracts, or will cure the symptoms when they have developed.

The anaemia which is a feature of the cobalt-deficiency syndrome in the ruminant is apparently the result of a breakdown of a link in the chain of metabolic events responsible for the maturation of the red cell, other than the one that fails in human pernicious anaemia.

In contrast to the dramatic change in the general health and activity of a deficient sheep which supervenes immediately on treatment with cobalt, the recession of the anaemia is slow, and so the low oxygen-carrying capacity of the blood *per se* is not the prime cause of the physiological disability but a symptom of a more general metabolic defect. *It appears that we may seek with confidence a less restricted physio-logical role for cobalt than the specific part it plays in the maturation of the red cells.*

And that is the present state of our knowledge of the function of cobalt in the ruminant. Our curiosity is now directed to the possible significance of cobalt in the nutrition of the associative growth of micro-organisms which normally inhabit the rumen. The A.P.A. factor is known to be necessary for the growth of the Dorner strain of *Lactobacillus lactis* and this implies that cobalt may be an essential adjuvant for other bacteria—the quantity required, $1.3 \times 10^{-5} \mu g$. of the crystalline factor (equivalent to $5 \times 10^{-7} \mu g$. Co) per ml. of medium for half maximum growth, suggests that the A.P.A. factor is among the most potent of the compounds essential for the proliferation of micro-organisms.
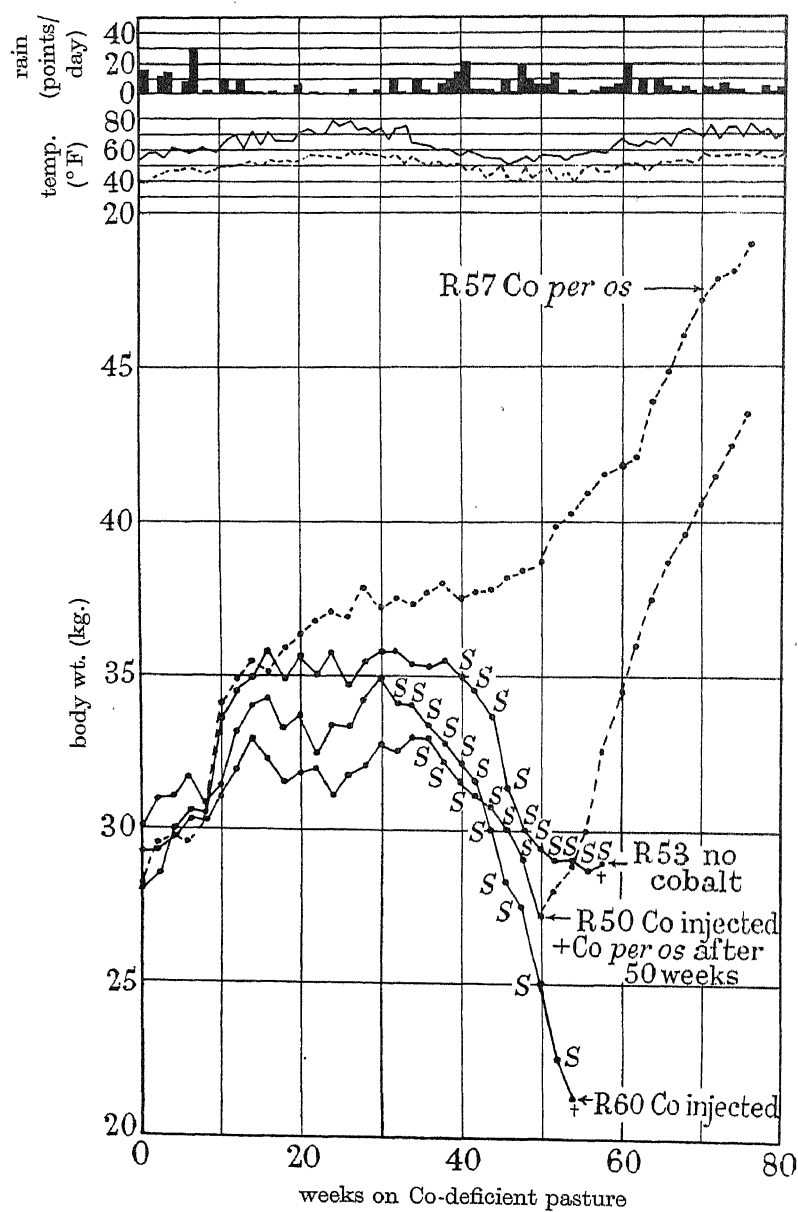
FIGURE 5. The individuals of this group of 4 Merino ewes (*aet.* approx. 9 months at beginning of the experiment) were grazed together on cobalt-deficient pastures and treated as follows: R. 57, with the equivalent of 1 mg. Co/day administered once a week *per os*; R. 60, and R. 50, with 1 mg. Co/day injected once a week into the jugular vein; R. 53 received no supplement of cobalt. All received adequate copper. After 50 weeks, when R. 50 was in the terminal stages of the deficiency syndrome, the equivalent of 1 mg. Co/day was superimposed on the injections. *S* denotes recognizable deficiency symptoms. The behaviour was similar to that of the animals in other groups. The cobalt concentration of the pastures varied between 0·02 and 0·03 μg. Co/g. dry wt. (From experiments of Marston & Lee 1943–44.)

Copper deficiency

It is possible to be more concise and somewhat more final in the discussion of our experience with copper deficiency—for reasons which may later become clear, problems associated with the physiological role of copper were pursued with more vigour and they have yielded more readily than those that concern the function of cobalt.

When the cobalt requirements of sheep depastured on the deficient littoral are fulfilled, a syndrome slowly develops which has proved to be the result of a serious copper deficiency; additional copper leads rapidly to its regression, and will maintain the animals in normal health and productivity.

Sheep are unable to establish a positive copper balance on these deficient pastures which contain, according to the stage of growth, between 2 and 4 μg. Cu/g. dry wt. Their reserves of copper are depleted and the reduced status is reflected by a steady fall in the blood-copper to levels below one-fifth of the normal concentration of approximately 1 μg./ml. They sicken and exhibit well-defined symptoms. By appropriate treatment with supplements of copper insufficient to maintain the normal physiological requirements, we were able to produce an ordered series of deficiency states that ranged from the frank syndrome in the untreated animals to an incipient state of deficiency in which overall efficiency was impaired without the appearance of specific symptoms (figure 6).

It was our experience with this series of experiments which led us later to recognize the aetiology of a number of ill-defined maladies of young grazing stock, enzootic over extensive areas distinct from the very deficient terrain upon which these observations were made. And they provided the second, and this time, much more definite clue to the understanding of the factors which had rendered agriculture impossible on the vast areas of the so-called mallee deserts.

The observations showed clearly that when the copper status of the sheep falls below a well-defined level, the first metabolic process to become seriously impaired is the chemical mechanism concerned with keratinization. Further depletion results in a disordered iron metabolism which becomes manifest in a hypochromic anaemia and the deposition of haemosiderin in the tissues. A very low copper status in the ewe during pregnancy leads invariably to nervous sequelae in her lamb.

Nervous sequelae of copper deficiency

The results of this nervous disorder have been known to shepherds for many centuries in the British Isles and elsewhere.

The symptoms range, according to the severity of the lesion, from complete paralysis at birth to a bland ataxia which may appear at any age up to 6 months or more. The incoordination of the hind limbs in the milder cases presents a clinical picture which suggests failure of the upper motor neurones; and histological examination of the cord reveals extensive demyelination in the ventro-lateral column always involving the direct and crossed pyramidal tracts which are not as highly differentiated in the sheep as in the primates.

The invariable nature and restricted distribution of these lesions raises many
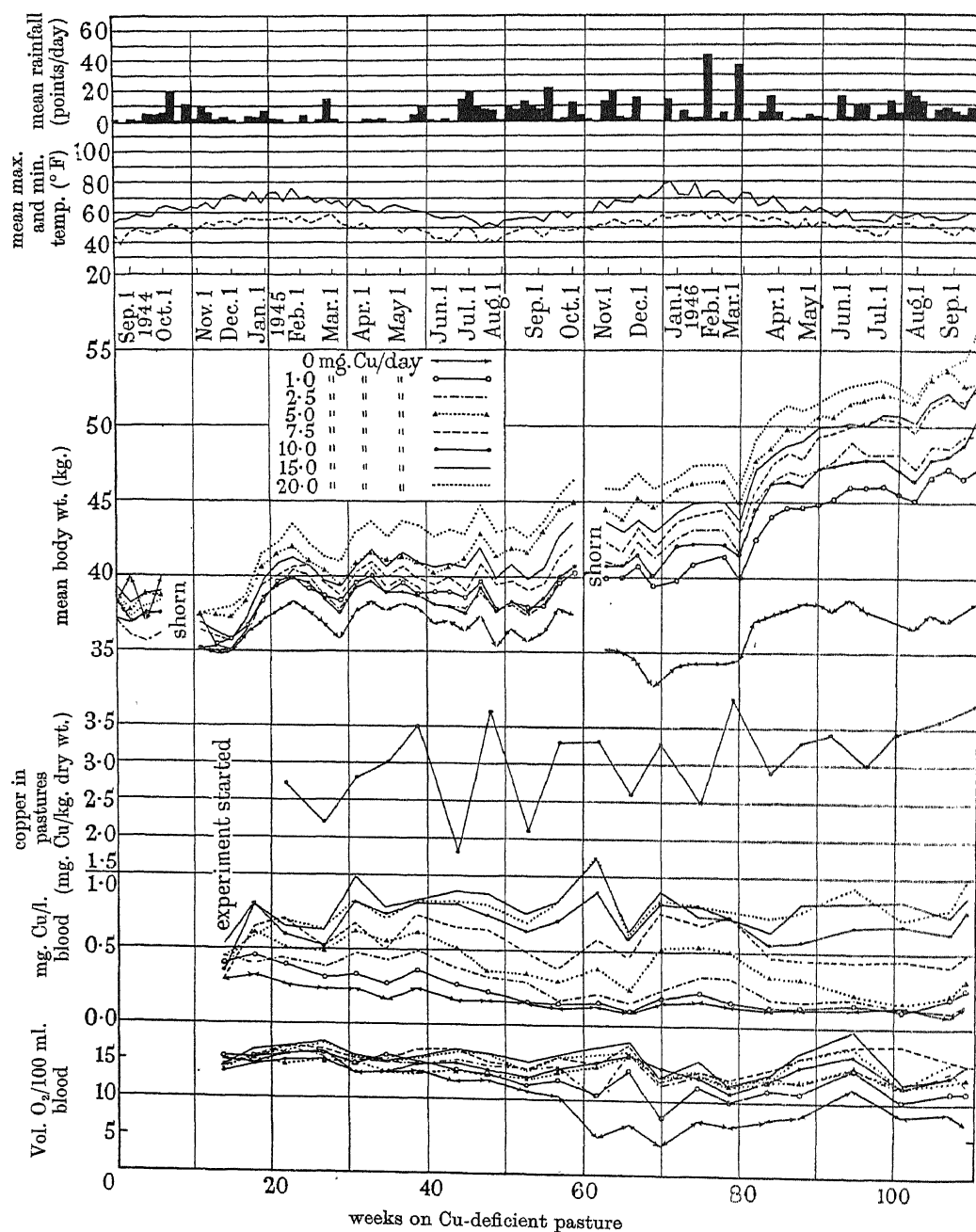
FIGURE 6. The data are from an experiment designed primarily to study the influence of various degrees of copper deficiency on the quantity and nature of the wool produced by Merino sheep. This flock was selected from a large number of young Merino ewes to be of even wool type. It was divided into 8 groups each of 8 animals and grazed on seriously copper-deficient pastures. All received cobalt. The influence of the supplements of copper on the oxygen-carrying capacity and on the copper concentration of the bloods is shown. The effects on the wool fleece are indicated in figure 8 and plate 4. (From experiments of Marston & Lee 1945-6.)

intriguing questions, but it is the effects of copper deficiency on the process of keratinization that I should like to relate in some detail to you.

Copper deficiency and keratinization

Copper deficiency reveals itself first as a marked deterioration in the quality of the fleece (plate 3*a*). A gradual series of morphological changes reflects in the wool staple the progressive decrease in the efficiency of chemical mechanisms in the follicles which require copper if they are to function normally. As the copper reserves are depleted the crimp becomes progressively less distinct until the fibres emerge entirely devoid of character.

The capacity of the follicles to impart crimp returns immediately on the resumption of a normal copper status and the staple that grows subsequently resumes boldly the character of the particular wool type of the individual (plate 3*b*).

There is thus little doubt that the changes in structure reflected by this lesion in the fleece are due to the breakdown of a catalytic process in which copper is primarily involved. The follicles of a black sheep seriously depleted of copper, lose their ability to impart crimp and their capacity to produce pigment. The role assumed by copper in effecting oxygen transfer is beautifully illustrated by the sudden reappearance of both melanin and crimp when the normal copper concentration in the follicles is reinstated (plate 3*c*).

Consideration of the cytology of the follicle leads to the site of this failure. As the proliferating cells at the base of the follicle mature, their expanding protoplasm impels them to present a minimum surface and fill with no interstices the restricted space they occupy so that each cell becomes faceted and approximates to the shape of Kelvin's minimum tetrakaidecahedron. In this state they are forced towards the conical constriction where they are gradually compressed to about one-eighth of their original cross-section and so become greatly elongated into spindle-shaped cortical cells which, when keratinized, comprise the wool fibre. During their extrusion through the constriction the globular protein of these cells granulates and proceeds through microfibrillae of fibrous protein (Vörner's trichohyalin of classical histology) to hard keratin.

A specific histochemical reaction with alkaline nitroprusside which forms a deep purple complex with thiol (—SH) groups reveals the process which is impaired by copper deficiency. When this reagent is applied to a normal fibre freshly plucked so as to retain its root, the basal cells are stained very lightly if at all. At the level where the intracellular structures begin to granulate and form into microfibrillae, the staining is intense and from there, in a normal fibre, it extends for approximately $100\,\mu$ to cease quite abruptly in the fully keratinized portion. When applied similarly to a wool fibre from a sheep with a low copper status the intense reaction begins at the same level but it extends for $1000\,\mu$ or more—practically the whole of the fibre embedded in the skin reacts strongly, the staining of the upper portion becoming progressively less intense as it approaches the skin surface (figure 7). On dosing the deficient animal with copper this greatly increased zone of free thiol groups retracts to its normal position within a few hours.

The evidence leaves little doubt that copper is primarily responsible for catalyzing the oxidative closure of the thiol residues of the prekeratin fibrous protein to the disulphide linkages of keratin. The rate at which this reaction proceeds is a measure of the copper status of the sheep. In the normal integument the oxidation of the reduced fibrous protein is complete in 8 to 12 hr.—the fibre is set deep down in the follicle: in the integument of a seriously copper-deficient animal this oxidative change takes 3 days or more to proceed to completion and so the fibre retains its plastic condition until it emerges at the skin surface.
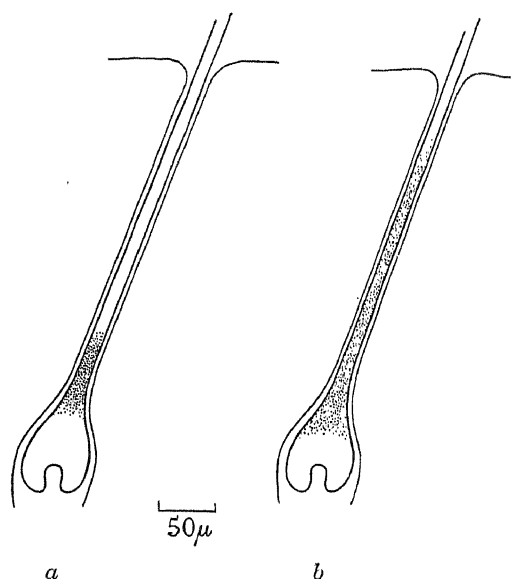
FIGURE 7. Wool follicles (diagrammatic) showing the approximate distribution of free —SH groups in the wool fibre during keratinization as indicated by the nitroprusside reaction applied to wool from a normal and b copper-deficient sheep.

The preferred molecular orientation in the wool fibre originates during the passage of the cells through the constriction at the base of the follicle—the normal fibre becomes anisotropic during its extrusion through this constriction and the molecular asymmetry remains. The long-chain fibrous protein molecules are alined while still in a reduced state and fixed in this position by copper-catalyzed oxidative closure of the cysteine residues of adjacent molecules to disulphide bonds which bind the keratin fibrillae together. In the copper-deficient sheep the comparatively slow oxidation of the fibre provides ample opportunity for subsequent disorientation before internal bonding is completed.

The aberrant physical properties of copper-deficient wool—its reduced tensile strength, its lack of elasticity and abnormal stress strain cycle, its tendency to become permanently set, its reduced affinity for dyes, etc.—with little doubt arise from the abnormal internal bonding and less orderly structure within the fibres. There is no gross change in its amino-acid constitution.

When the copper status of a sheep falls below normal, the efficiency of inter-mediary metabolism is impaired; appetite is lessened and so the rate of wool growth

a Typical calcareous dunes

b Typical mallee heath

Areas of aeolean soils in Southern Australia where trace element deficiencies prevail.
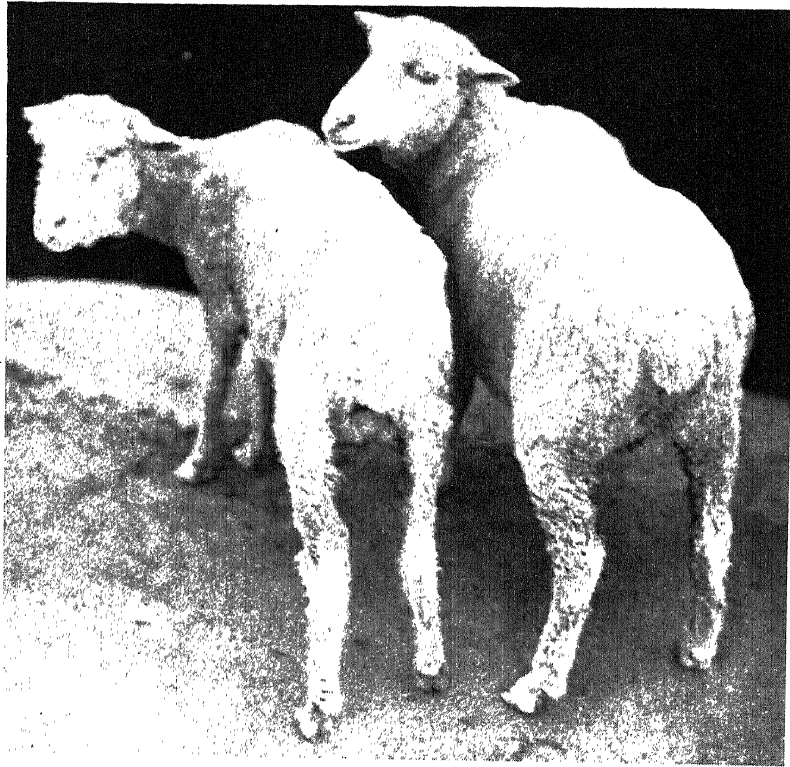
a
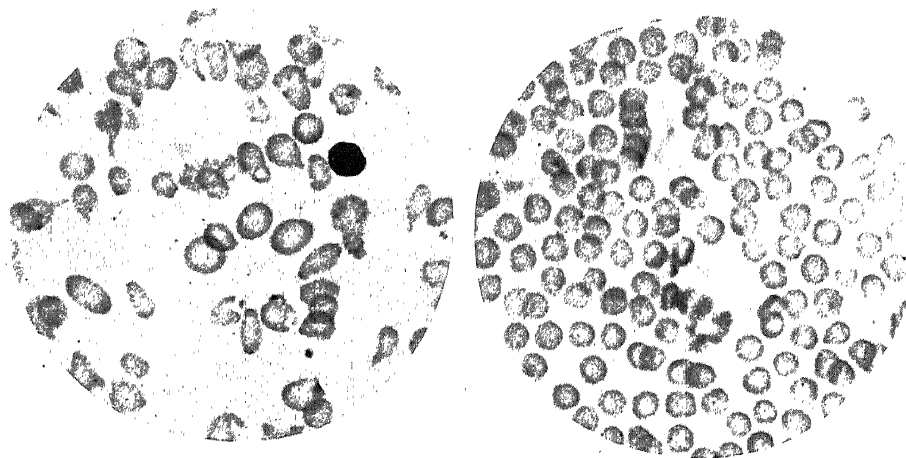
b

The effect of cobalt deficiency on Merino sheep.
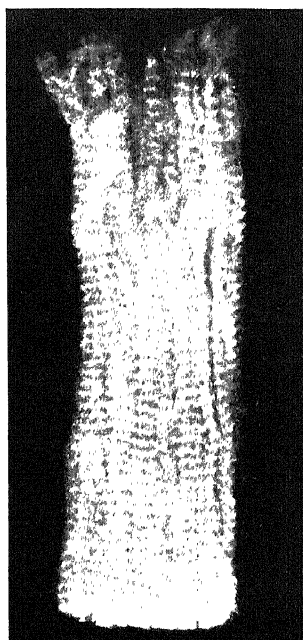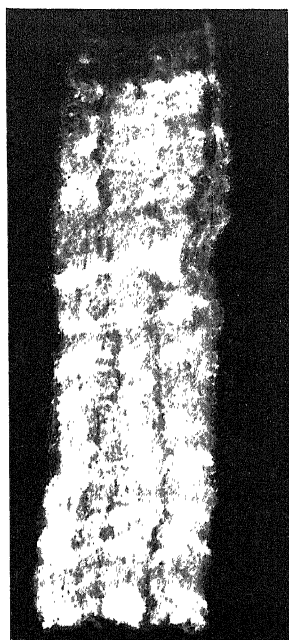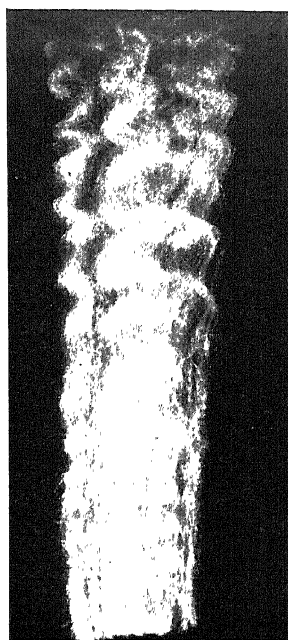
1941　　　　　　　　1942　　　　　　　　1943

a
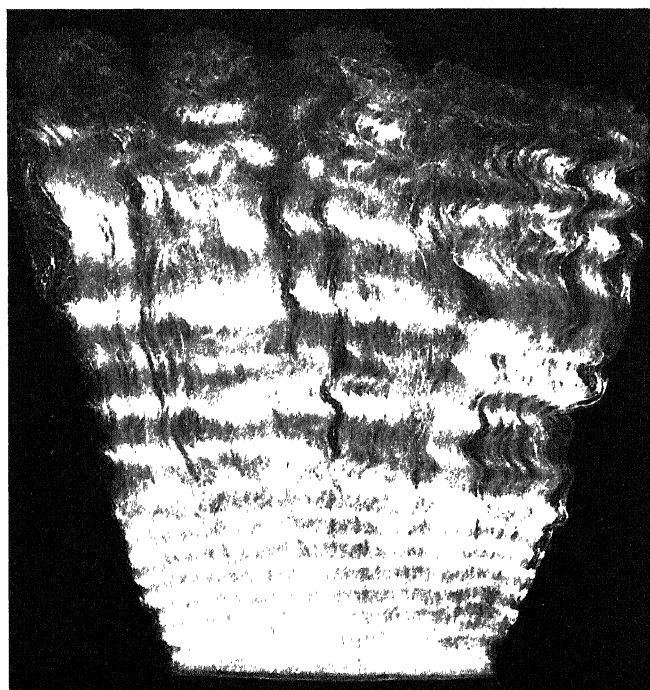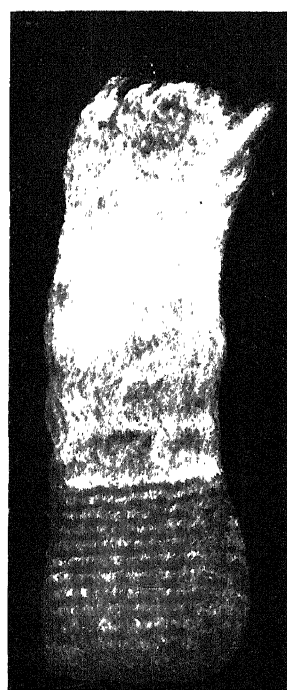
b　　　　　　　　　　　　　　　*c*

The effect of copper deficiency on the wool grown by Merino sheep.
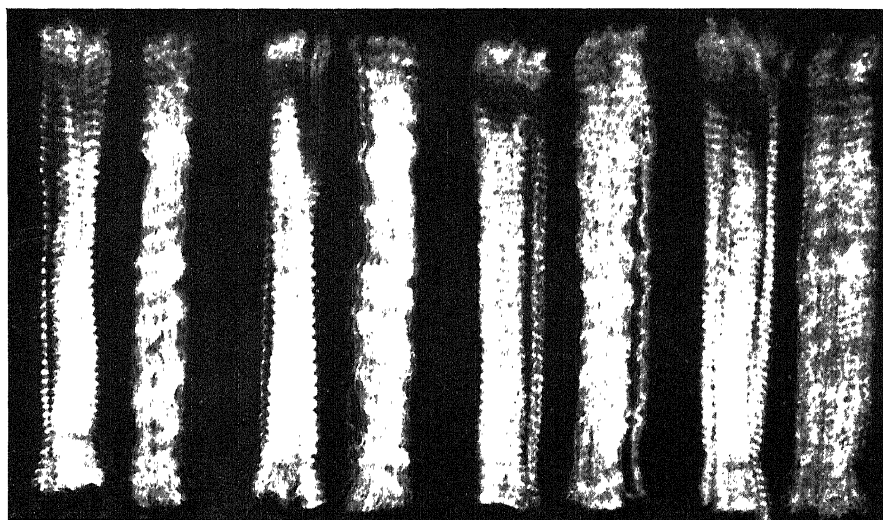
1944   1945   1944   1945   1944   1945   1944   1945
  no Cu   1 mg. Cu/day   2·5 mg. Cu/day   5·0 mg. Cu/day
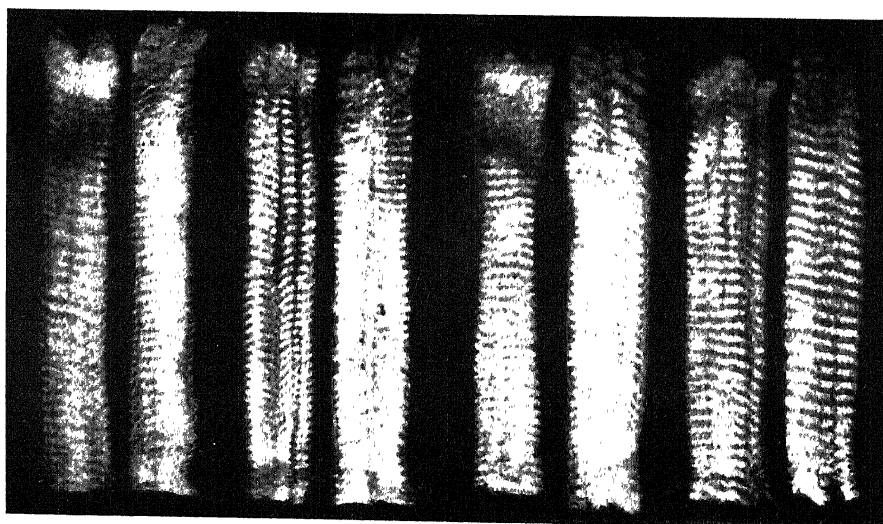
1944   1945   1944   1945   1944   1945   1944   1945
 7·5 mg. Cu/day   10 mg. Cu/day   15 mg. Cu/day   20 mg. Cu/day

The effect of different degrees of copper deficiency on the character of Merino wool.

a

b

Resistance of certain plant species to soil conditions which impose copper
deficiency on other species.

a
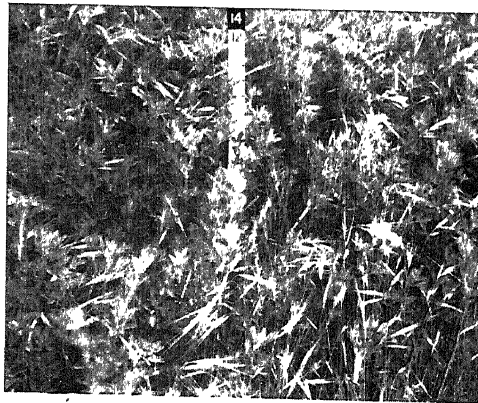
b   *c*

The establishment of lucerne on the calcareous littoral.
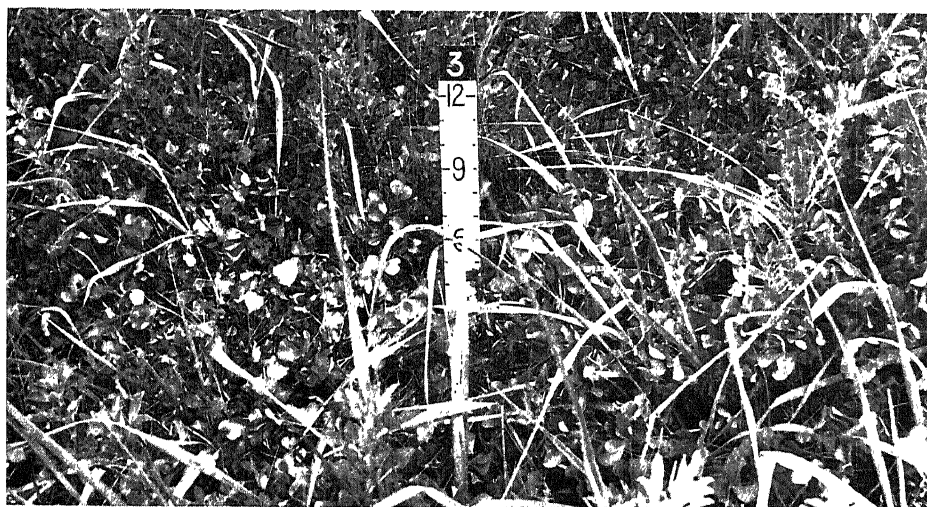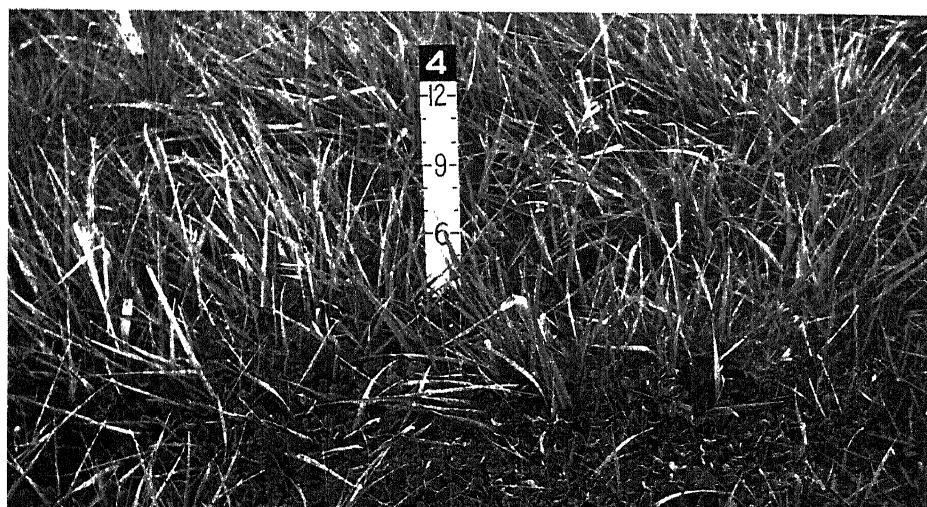
The development of pastures on the mallee heath soils of the hinterland,
S.E. South Australia.

is diminished. We employed methods analogous to a chemical titration to estimate the extent of this. Supplements of copper ranging in increments from nil to amounts that were more than sufficient to meet the nutritional requirements were administered to a series of carefully matched groups of sheep depastured as a flock on the frankly deficient terrain (figure 6). The mean weight of clean-scoured wool produced was found to vary exponentially with the amount of copper provided in the supplement. In the nutritional environment of the littoral dunes, the asymptote of maximum response was reached when between 7·5 and 10 mg. Cu/day was supplied. Equilibrium was established at this level of intake and greater amounts of copper brought about no further improvement either in the quality or in the quantity of wool produced (plate 4 and figure 8).
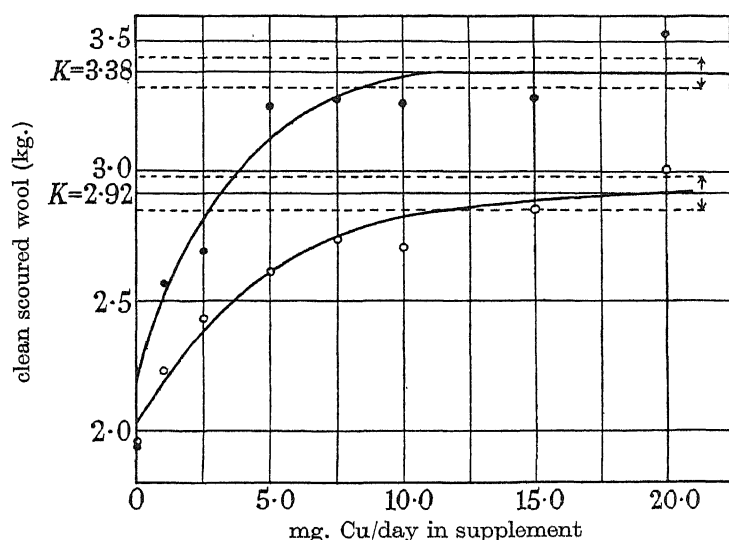
FIGURE 8. The influence of a graded series of copper deficiency (see figure 3 and plate 4) on the annual production of clean scoured wool during the first 2 years of the experiment is shown (O 1945 clip, ● 1946 clip). The observed means for each group are plotted in relation to curves derived by the method of least squares to specify that the increase in wool production per unit of copper supplement is proportional to the deficit below the maximum wool production K. The asymptote K is indicated with its s.D. plotted on either side. In each year the maximum response was achieved when the supplement provided between 7·5 and 10 mg. Cu/day. The pastures were more nutritious in 1946 than in 1945.

Adequate copper more than doubled the monetary value of the fleeces from sheep on this very deficient terrain and the fleeces from the deficient groups ranged in value according to the degree of deficiency imposed. The reduced value of the fleece from a copper-deficient sheep is not due entirely to its lighter weight. The lack of crimp is recognized by wool-buying appraisers to indicate properties that have been found undesirable in technological processing and so copper-deficient wool is heavily discounted.

These titration experiments allowed the consequences of all degrees of copper deficiency to be expressed in language readily appreciated by pastoralists and general adoption of preventative measures followed.

Delineation of deficient areas

In the preliminary investigations, geochemistry with its full appreciation of the petrological aspects of soil origin provided valuable guidance, and in virgin country ecological studies have been helpful in delineating deficient areas within any one climatic zone.

On the very deficient tracts where only that vegetation inured by selection to the inadequate nutritional environment is able to persist, the small quantities of micro-nutrients which become concentrated in the fodder plants are grossly insufficient to fulfil the requirements of grazing ruminants.

On less seriously deficient terrain, however, the chemical composition of the available pastures is subject to a great variety of environmental influences: those of climate—the seasonal effects of light, rainfall, temperature, etc.; those of the soils —the nature and interaction of the minerals of the soil proper and of the parent rocks, the pH, the structure, the microflora, etc.; and those of the vegetation itself —the species, variety, strain of the plants, their root systems, their stage of growth, the mineral interactions within them, to mention but a few, are all reflected in the composition of the fodder. And so it is not feasible, under these conditions, to assess the status of micro-nutrient elements without some means of integrating this complex and labile system of variables. Indicator plants do this, and if appropriate species are selected, specific lesions are exhibited. The grazing animal further resolves these factors and their influence may be measured in terms of animal production.

We have exploited specific symptoms such as the copper-deficient lesion in the wool for mapping the areas that are seriously involved and for revealing incipient deficiencies of copper on terrain where hitherto such limitations had never been suspected.

The extent of the areas where less obvious deficiencies are prevalent is very much greater than that of the frankly deficient terrain, and the consequences are of little less economic significance. The recognition and correction of deficiencies on these great tracts of marginal lands followed experiments with indicator plants and led finally to the development of pastures on the very deficient aeolian soils.

Solution of the problems of the aeolian soils

Development of pastures on the littoral

From the beginning the factors which limited so strikingly the development of vegetation on these huge tracts intrigued us.

In our preliminary experiments with animals grazing on the littoral we attempted to establish a sown pasture of a high-yielding perennial grass and a legume suited to this climatic zone by applying liberal dressings of phosphate, nitrogen and potash. The manurial treatment served only to stimulate the grasses already prevalent— the sown species failed completely (plate 5b).

Subsequently, when we realized that copper deficiency in the fodder limited the welfare of animals depastured there, we were led to investigate the effects of increasing the level of available copper in the soils. It was a series of experiments with oats

as an indicator plant that proved the low fertility of these calcareous sands to be due primarily to a low copper status.

The experience of early settlers who had attempted without success to farm these deficient tracts suggested that rye would grow there but oats, wheat and barley would fail. This was substantiated. Oats proved particularly susceptible, for when grown in this environment the oat plant displayed flagrant lesions and failed to head (plate 5a). Plot experiments of factorial design revealed that a dressing of 7 lb. $CuSO_4.5H_2O$/acre applied at sowing prevented these deficiency lesions and a crop of 23 bushels/acre resulted—nitrogen was then the limiting factor for when nitrogenous manures were superimposed the yield was more than doubled. An identical dressing of copper applied a year before sowing resulted, under the same conditions, in a crop of 62 bushels/acre. Here is a probable effect of copper on nitrification within the soils.

In the absence of additional copper all legumes failed completely, but with it lucerne (plate 6), black medic, barrel medic and other pasture legumes developed with surprising vigour, and under these conditions pasture grasses grew well in association.

The effects of the copper dressings proved lasting—splendid high-yielding mixed pastures established in this way several years ago have persisted without further additions of copper—the yield tending to increase rather than diminish with time.

The factors which had rendered agriculture impossible on about 2000 square miles of littoral were thus revealed and application followed.

Development of pastures on the siliceous hinterland

Although experimental flocks confined to the great hinterland of siliceous sands provided evidence of serious copper deficiency in the sparse herbage on which they subsisted, our experiments there with oats as an indicator proved zinc and not copper to be the first limiting factor.

When the phosphate level of these virgin soils was raised, oats grew and yielded a poor crop of less than 10 bushels/acre. No apparent lesion of copper deficiency appeared and dressings of copper produced no significant response in the yield. But very striking symptoms of disordered metabolism became evident—the older leaves assumed a deep purple colour. A dressing of 7 lb. $ZnSO_4.6H_2O$/acre prevented the accumulation of anthocyanin and doubled the yield before the low nitrogen status of these soils became the limiting factor. Lucerne responded to copper but not to zinc—here again a species difference. Subterranean clover failed completely to develop without additional zinc, grew luxuriantly with it, but in the absence of a dressing of copper failed to set viable seed and so did not persist.

Further series of experiments proved unequivocally that splendid, high-yielding permanent pastures could be established by sowing the perennial grass, *Phalaris tuberosa*, and the annual, subterranean clover, with a dressing of superphosphate enriched with 7 lb. each of $ZnSO_4.6H_2O$ and $CuSO_4.5H_2O$/acre (plate 7), and that —an important economic consideration this—these pastures could be established under a cover crop of wheat or oats, the value of the grain from which would more than discharge all costs of the sowing.

The main factors which have limited the development of vegetation on these huge areas of hungry soils have been revealed, the course that agricultural development should take is clear, and already application is proceeding apace, with dramatic results. About 3000 square miles of sandy heath is involved.

CONCLUDING REMARKS

Although application of these findings met with spectacular success and the implications are of considerable economic importance, there is no need here to confess that application was not the immediate goal for those of us who carried out the experiments. Our delight has been more with the phenomena that have been revealed, with the problems that these have clarified, and with the pattern as it emerged. Each step allowed a glimpse of physiological mechanisms in which minute traces of heavy metals play an essential part. The urge to add further to this picture will continue to be our main stimulus: I have no doubt that pursuit of knowledge along these lines will pay economic dividends.

The realization that research in the applied field may miss the substance for the shadow if it becomes too practical, was the central thought in the minds of the small group who launched C.S.I.R. on its course of free scientific inquiry, and who had the foresight in the beginning to ensure that official science in Australia would be free from political control, and independent of public service administration.

All of us who have worked in the splendid atmosphere which this created realize the wisdom and vision of the three men who were mainly responsible—a statesman, Viscount Bruce, F.R.S., who as Mr Stanley Bruce was the Prime Minister of Australia 22 years ago when C.S.I.R. was founded; an engineer, Sir George Julius, our first Chairman, who died in office; and a chemist, Sir David Rivett, F.R.S., who left the Melbourne Chair to set the course and take the helm, and who has since tirelessly guided C.S.I.R. on the not particularly easy route of free scientific inquiry.

Our gratitude to these men, and our affection and respect for our leader, are unlimited. We feel that we are the subjects of a grand experiment in official scientific administration and this, perhaps more than anything else, stimulates us to carry the findings of our scientific inquiries a good deal further into the field of practical application than we should be inclined to under any other regime. We are all anxious to prove this experiment a success.

APPENDIX ON HISTORY AND ORGANIZATION OF DIVISION OF BIOCHEMISTRY AND GENERAL NUTRITION

The nucleus of the Division of Biochemistry and General Nutrition was formed when the Council for Scientific and Industrial Research founded the Division of Animal Nutrition in 1927 with the appointment of Professor Thorburn Brailsford Robertson as Chief. Robertson built the central laboratory in the grounds of the University of Adelaide, drew together a small team, and with courage and vision set to work on problems associated with the nutrition of sheep.

Shortly after Robertson died in 1929, Sir Charles Martin, F.R.S., then recently retired from the Directorship of the Lister Institute, London, showed his affection

for Australia by accepting the responsibility of guiding the young Division for a short period. He was Chief of the Division of Animal Nutrition during 1930–32. The wise and inspiring leadership of these men initiated a spirit of research that has held the team together and has directed their efforts. Subsequent to the return of Sir Charles Martin to England in 1932 the Division of Animal Nutrition became part of the newly created Division of Animal Health and Nutrition with Dr L. B. Bull as Chief. In 1944, to simplify administration, the research activities centred at Adelaide were raised to the status of a Division of the Council and the scope was increased by the reference implied in the title—Biochemistry and General Nutrition.

During these changes the course of the research has altered very little—the Division has gradually consolidated into a Research Institute which is concerned broadly with the physiology and biochemistry of ruminants and more particularly with the nutrition of the wool sheep.

The original two-storied Animal Nutrition Laboratory, its annexes which house the two calorimeters, metabolism cages for 20 sheep and subsidiary apparatus, and the four-storied Biochemistry Laboratory now nearing completion which comprise the buildings where the Division is centred, are situated within the grounds of the University of Adelaide. This proximity has fostered a close liaison between the Division and the University; the research staff share alike the amenities and the disciplines of their university colleagues. The Division has a central field station, Glenthorne, of some 600 acres situated about 11 miles from the laboratories, and a series of field stations distributed over the pastoral areas of Australia at sites where nutritional disabilities exist. Experiments conducted on these outlying field stations serve both to clarify the problems and to apply under conditions of station practice the information which results from more intensive studies in the laboratory.

The Division is organized into a number of sections which may be defined loosely as nutritional biochemistry, vital energetics, tissue metabolism, microbiology, analytical and physical chemistry, organic chemistry, animal physiology, plant nutrition, and field stations. The sections flow imperceptibly into one another, the fluidity being fostered to ensure the unhampered intercommunication of ideas essential for successful team work. Long-term individual and collaborative studies which aim at extending knowledge of the nutritional physiology of the sheep constitute the main experimental projects, but on occasions, relatively short-term experimental investigations are undertaken to illuminate the origin and nature of specific nutritional disabilities which climate and terrain impose on grazing ruminants. The guiding principle of all investigations is the simple faith that practical application will flow naturally and freely from a more complete understanding of the underlying phenomena.

The reference of the Division is to conduct research; the research findings are extended to practice in the pastoral industry through the appropriate organizations with which there is a close liaison.

T

DESCRIPTION OF PLATES 1 TO 7

Plate 1

Typical areas of aeolian soils in Southern Australia where trace element deficiencies prevail

a The Recent calcareous dunes of the littoral consist essentially of comminuted shell fragments which overlie unconformably and at considerable depth the older soils. The unusually small amount of copper available from these soils limits the development of most plant species. These deficient tracts are overgrown with the calciphile grasses *Bromus madritensis* and *Lagurus ovatus*. Fodder plants which grow there in the absence of manurial dressings of the elements which are short in these soils do not provide enough cobalt and copper to fulfil the physiological requirements of grazing ruminants.

b The siliceous, solonetz soils of the hinterland support a stunted heath inured by natural selection to this singularly deficient terrain. Growth of other plant species is limited by the deficiencies of phosphorus, zinc and copper which prevail. Ruminants depastured on these tracts develop symptoms of copper deficiency.

Plate 2

The effect of cobalt deficiency on Merino sheep

a The twin Merino ewes are those referred to in figure 2. The photograph was taken in December 1938 after they had been depastured on cobalt-deficient terrain (see plate 12 a) for 55 weeks during which the ewe on the right, W 6.248, had received the equivalent of 1 mg. Co/day administered *per os* thrice weekly. The other, W 6.241, which was untreated during this period exhibits the symptoms of extreme cobalt deficiency. The oxygen-combining capacity of the blood of the former was at this time 15·5 vol. O_2/100 ml.; that of the latter was 3·5 vol. O_2/100 ml.

b The morphology of the formed elements in the blood of the sheep above is shown. The film from W 6.241 is typical of the blood dyscrasia which supervenes in the terminal stages of cobalt deficiency; that from W 6.248 is normal.

Plate 3

The effect of copper deficiency on the wool grown by Merino sheep

a The capacity to impart the crimp which is a feature of Merino wool is lost by the follicles when the animal's copper status is seriously depleted. The three staples illustrated were drawn in three successive years from an area defined on the shoulder of an experimental ewe; the first was grown while the animal was on normal pastures, the other two reflect the deterioration of the animal's copper reserves during two years' grazing on copper-deficient terrain.

b The ability to impart crimp to the fibre returns abruptly when the copper necessary for this function is provided. The staple illustrated was taken from an experimental ewe which had been dosed with the equivalent of 10 mg. Cu/day while grazing on the deficient pastures on which it had been depleted.

c When copper-deficient, black-woolled sheep lose their ability to produce melanin and their capacity to impart crimp to their wool fleece. Both functions return immediately when a normal copper status is resumed. The effect of a supplement of copper on the pigmentation and crimp is illustrated. (From experiments of Marston & Lee.)

Plate 4

The effect of different degrees of copper deficiency on the character of Merino wool

When a sheep is confined to pastures which are acutely copper-deficient its reserves of copper are depleted rapidly. The rate of this depletion may be decreased by providing amounts of copper which are insufficient to reinstate a positive copper balance. In this way an ordered series of deficiency states may be imposed. The series of staples shown are from ewes em-

ployed in an experiment (figures 6 and 8) which aimed, by means of a titration, to determine the quantity of copper necessary to reinstate copper equilibrium in sheep depastured on copper-deficient terrain. Each pair of staples was drawn from an identical site in two consecutive years, the first of each were grown while the animal was on normal pastures, the second during the first year of treatment while on deficient pastures. The deficiency lesion appeared in all staples from sheep that received less than the equivalent of 7·5 mg. Cu/day in the supplement. These form a graded series. (From experiments of Marston & Lee.)

Plate 5

Resistance of certain plant species to soil conditions of the calcareous
littoral which impose copper deficiency on other species

a Experimental cereal crops sown with superphosphate but without copper. Rye corn growing normally under conditions in which oats (var. *mulga*) have developed flagrant lesions of copper deficiency and have failed to head.

b Complete failure of certain species of pasture plants (*Phalaris tuberosa*, *Trifolium subterraneum* and *Medicago lupulina*) sown on the calcareous littoral sand manured with ample dressings of superphosphate, potassium sulphate and ammonium sulphate. Left: The sown species have failed to develop and the area is overgrown with the natural species *Lagurus ovatus* and *Bromus madritensis* which have responded to the better nutritional environment provided by the dressings. Right: The natural cover unmanured. (From experiments of D. S. Riceman.)

Plate 6

The establishment of lucerne on the calcareous littoral

a Typical shell-sand dunes at Robe, South Australia.

> *Top left:* Area outside rabbit-proof fence denuded of grass-cover by rabbits.

> *Top right:* Typical natural pasture of *Lagurus ovatus* and *Bromus madritensis*.

> *Bottom left:* An experimental crop of lucerne (*Medicago sativa*) established with manurial dressings of copper sulphate.

> *Bottom right:* Experimental quadrates showing effect of the application of copper sulphate.

b Quadrate (on bottom right area in *a*) sown May 1938 with lucerne together with manurial dressing of superphosphate and copper sulphate. The lucerne is established. Photographed November 1939.

c Quadrate adjacent to *b* sown May 1938 with lucerne together with manurial dressing of superphosphate. In the absence of additional copper the lucerne has failed and the area is dominated by *Lagurus ovatus* and *Bromus madritensis*. Photographed November 1939. (From experiments of D. S. Riceman.)

Plate 7

The development of pastures on the mallee heath soils of the hinterland S.E. South Australia

a Area cleared for experimental trials.

b and *c*. Experimental pasture of *Trifolium subterraneum*, *Phalaris tuberosa* and some *Medicago sativa* developed on mallee heath soils with manurial dressings of superphosphate, zinc sulphate and copper sulphate, *b* in late spring of second year and *c* in early spring of fifth year after sowing. (From experiments of D. S. Riceman.)

A discussion on muscular contraction and relaxation: their physical and chemical basis

INTRODUCTION

The elementary unit of contraction in skeletal muscle is a single twitch, evoked either by a single nerve impulse arriving at an end-plate or by a direct electric shock. It is a very rapid affair, requiring for completion a time varying from a few milliseconds to a fraction of a minute according to the muscle and its temperature. It is followed by a slow recovery ('recharging') process taking from a few minutes to an hour.

The focus of the present discussion is the relation between (*a*) known mechanical, thermal, electrical and other physical changes accompanying contraction and relaxation and (*b*) the underlying chemical causes. The fundamental difficulty is that normal chemical methods are very slow and insensitive in face of the requirements posed by the muscle twitch. To illustrate this, the heat production in a twitch, say 3 mcal./g., if derived (for example) from the breakdown of creatine phosphate, would require only $2 \cdot 5 \times 10^{-7}$ g.mol. of phosphate, per gram of muscle, to be set free. To detect such a quantity at all is a formidable enough task; to resolve its appearance in times of the order of tenths or hundredths of a second is one of fantastic difficulty. Even in the very slow twitch of a toad's muscle at 0° C the state of full activity is reached within 30 or 40 msec. after a stimulus, and relaxation is already obvious at 0·3 or 0·4 sec.

It might be thought that the chemical problem could be eased by using a succession of stimuli and a maintained contraction, thereby increasing both the times and the quantities involved. As regards the slow processes of recovery, within limits that is true; in relation, however, to the rapid initial processes it is sometimes apt to be highly misleading. If the whole cycle of contraction and relaxation is com-

[40]

pleted, in a frog's muscle at room temperature, in 0·1 sec., the resolution of chemical events within that cycle is merely confused by repeating it (say) thirty times a second for several seconds.

There is no direct way of mitigating this difficulty except to improve the sensitivity and speed of chemical methods and to work with the slowest possible muscles at the lowest possible temperature. Lacking direct chemical evidence physical methods must be used. Many of these are very sensitive and rapid; others could certainly be improved in speed and sensitivity if needed. The difficulty is to relate their results specifically to chemical processes. To suggest means of overcoming that difficulty is the primary object of this meeting.

Following is a summary of some of the physical information now available.

Mechanical

The contractile component of a muscle is in series with an undamped elastic component. In parallel with both (in skeletal muscle) are other elastic components; the complication introduced by the latter can be largely avoided by working at lengths not greater than the resting length in the body, the parallel elastic components then become relatively slack. All the elastic components show the usual load-extension relation of biological fibres, being much more compliant at small loads than at greater ones. At the full isometric tension of a tetanic contraction the contractile component has shortened, and the series elastic component has been stretched, by 4% or more—depending on the length of the tendon—of the length of the whole muscle.

At rest the contractile component is plastic and extensible. Released under zero load its elements, either structural or contractile, become slack, as is shown by the lengthening of the latent period when a first shock is applied; the contractile component has to shorten and take up the internal slack, or its molecular linkages have to be reset, before external movement can be manifested or external force exerted. A second or third shock, however, is followed by the usual short latent period, showing that the slack has been taken up. A toad's sartorius under zero load can be brought by a few shocks to a length about one-half of that which it occupied in the body, and there it remains. The latent period is then little more than at its normal length, showing that the slack, if any, is very small. From its shortened condition it can be gently pulled out to any intermediate length, and its latent period remains small. Pulled out to its full length and then released under zero load its latent period to a first shock is long again, but becomes short once more when a second shock is applied. Relaxation is not an active process; no lengthening occurs in relaxation unless a muscle is loaded. The process can be repeated indefinitely. The phenomenon is important in showing that the organized molecular state of readiness to contract when stimulated is similar at different lengths. It is obviously desirable to examine the birefringence, the X-ray diffraction and possibly the dielectric constant of resting muscle in the range of lengths at which it can exist under zero tension.

Experiments made with a piezo-electric method of high speed and extreme sensitivity, on a muscle under initial tension, have revealed a very small transitory

lengthening preceding the usual shortening; but in a muscle under very low tension this lengthening is absent. The heat production begins at a high rate before the mechanical response can be detected. Something has happened during the mechanical latent period, its nature at present unknown; it must include the propagated physico-chemical change which travels inwards from the surface where excitation occurs, to produce the abrupt change of state next described.

It was natural to suppose that the intensity of activity of a muscle when excited (a) follows the course of the myogram, rising gradually after the latent period, reaching a maximum and declining during relaxation, and (b) is greater in a tetanus than in a single twitch. Quick stretches, however, applied at various moments show that both conclusions are wrong. The intensity of activity in a contracting muscle, defined as the force it can just bear without lengthening, is greatest very early after a shock, and is as high in a single maximal twitch as in a maintained tetanus. There is an abrupt transition from rest to full activity at the end of the latent period. In a twitch the intensity of activity remains on a plateau for a time, then declines in the process recognized as relaxation; in a tetanus, each shock restores it to its full value. The reason why the tension in a twitch is usually so much less than in a tetanus is simply that in the twitch there is too little time, before relaxation sets in, for the contractile component to complete its extension of the elastic component.

At a low temperature, with a low enough frequency of stimulation, an isometric contraction can be maintained for a long time. If released, the muscle shortens at the same speed whether release be early or late. If the stimulus is stopped, the muscle relaxes more slowly the longer the preceding stimulus. The mechanical properties of the active state are best examined during a maintained contraction, where the complication due to relaxation is avoided, but there is every reason to regard them otherwise as the same in a twitch. They are fully defined by the characteristic relation between speed of shortening and load; or, strictly speaking, by a family of such curves at different lengths. The characteristic equation

$$(P+a)\, v = b(P_0 - P)$$

connecting P the load to v the velocity, has three constants, a with the dimensions of force, b with the dimensions of velocity, and P_0 the maximum force which the muscle can exert. P_0/a is rather constant, being usually about 4; b increases considerably with temperature and is widely different in different types of muscle, the maximum velocity of shortening (under zero load) being generally about $4b$. The constant a is of particular interest, because it appears that the 'heat of shortening' (see below) is ax, where x is the distance shortened. The characteristic equation is obeyed rather accurately by a variety of muscles (man, frog, toad, tortoise); but a discontinuity occurs at zero speed, the velocity of lengthening under a force greater than P_0 being considerably less than calculated from the equation. This is not yet understood.

Under a load rather greater than it can bear an active muscle lengthens slowly; under a considerably greater load it 'gives' or 'slips'. We can regard the first process as 'reversible' in the thermodynamic sense, the second as largely 'irreversible'.

The second process is analogous to the 'cold-drawing' of a metal wire, or of a thread of polythene or nylon. When the excessive load is removed the muscle begins to shorten again in the usual way. In normal relaxation under a load the same kind of slipping occurs and the mechanical energy of the load is turned into heat.

Thermal

Heat is produced in muscle activity in two main phases: (*a*) during the mechanical response, 'initial heat'; (*b*) after the mechanical response, 'recovery heat'. The recovery heat is very slow; presumably it represents the balance of the heats of reaction of all the processes, exothermic and endothermic, by which the muscle is restored, after activity, to its chemical *status quo*. In the presence of oxygen the total recovery heat is about equal to the total initial energy set free, as work and heat together. No certain evidence exists that any of the chemical changes at present known, or believed, to take place as the result of muscular activity, occur otherwise than in recovery. It is true that if a muscle is stimulated by a succession of shocks, chemical changes (lactic acid formation, phosphagen breakdown, oxygen consumption, alterations of pH, etc.) are found to occur; but these might be associated with recovery from the earlier elements of the response to stimulation, and there is nothing to show that any one of them takes place during a single twitch; indeed, it seems fairly certain that most of them come afterwards. There is no direct evidence that the breakdown of adenosine triphosphate, for example, ever occurs in living muscle except under conditions of extreme fatigue verging on rigor. Plausible hypotheses exist, based on the study of enzyme systems *in vitro*, to explain why no such breakdown is found; but these cannot be regarded as evidence that it actually occurs. Indirect evidence suggests that ATP occupies a key position in the chemical mechanism of contraction; but it remains possible that its intervention is confined to the chemical processes of recovery.

The initial heat in a single twitch (amounting usually to about 3 mcal./g.) is composed of two parts, 'heat of activation' and 'heat of shortening'. No heat at all is given out in relaxation, unless mechanical work (the energy of a load previously lifted, or the tension energy of the series elastic component previously stretched) is degraded into heat in drawing out the relaxing contractile material; the heat then appears and the work disappears at precisely the same time. At no stage in contraction or relaxation has an actual absorption of heat ever been found. If endothermic processes occur they are entirely masked by exothermic ones.

When a muscle shortens it gives out extra heat, in amount proportional to the shortening. Let the shortening be x cm.; then the heat of shortening, in mechanical units, is ax g.cm., where a is a constant of the order of 400 g. weight for a muscle of 1 cm.^{2} cross-section. It is obviously significant that a is the same quantity as was found from the form of the force-velocity relation, without any thermal measurements at all. The heat of activation begins at its maximum rate before the earliest detectable sign of the mechanical response. It continues at a decreasing rate, finally vanishing at about the moment when relaxation sets in. It is not affected by shortening or by the performance of work. It is little influenced, if at all, by the length of the muscle. It is presumably the waste heat of the chemical

reactions by which the transition from rest to activity occurs. In a prolonged isometric contraction, after internal shortening is complete, the so-called 'heat of maintenance' is the summed effect of the heats of activation of the processes by which the state of full activity is reinstated by each successive element of the stimulus. The heat of maintenance falls off gradually during continued stimulation, corresponding to the decreased speed of relaxation found if the stimulus is stopped; both finally reach a constant value. Like the heat of activation the heat of maintenance is little affected by the length of the muscle.

In contraction, in addition to heat of activation and heat of shortening, energy is given out as external mechanical work. We write

$$E = A + ax + W,$$

where E is energy, A is activation heat, x is shortening and W is work. E, A, x and W are all functions of the time, and the equation applies not only to the whole process but to any part of it. The work W can be varied by varying the load, and provided that the shortening x is not changed the heat is completely unaffected by varying the work from practically zero up to the maximum which the muscle can perform. This (as we shall see later) has an important bearing on the thermodynamic nature of the machine. Let P be the load; then the *rate* at which energy is being given out (in excess of the activation heat) is $(P+a) \, dx/dt$. It is found experimentally that this is a decreasing linear function of the load, and we write

$$(P+a) \, dx/dt = b(P_0 - P).$$

Here a, b and P_0 are the same quantities as appear in the characteristic equation relating speed to load, b being a velocity and P_0 the maximum force which the muscle can exert.

If we could regard the heat of shortening as work degraded into heat in the process by which the molecules of the contractile system are forced to change their relative positions in shortening, the statement embodied in this equation would become much simpler; the muscle would give out energy in two forms only, heat of activation and mechanical work. There are several objections to this view, but the most decisive is the fact that when a muscle is slowly stretched during continued stimulation the work done in stretching it disappears—that is to say, it cannot be accounted for either as extra heat produced, or as the mechanical potential energy of increased tension. Moreover, if a muscle is stretched during the rising phase of a single twitch, the total heat set free may be equal to, or less than, the work done in the stretch, the net heat produced by the muscle itself being zero or negative. Presumably the work which disappears is absorbed in driving backwards chemical reactions which previously occurred. The motor, in fact, is used as a dynamo when the tram runs down-hill. Moreover, recent experiments have provided strong evidence that the converse of the heat of shortening occurs—the heat of lengthening is negative. When a muscle shortens it gives out heat in proportion to the shortening; it does work in addition, equal to the product of the shortening and the load. When it lengthens the heat of lengthening is negative and work is absorbed. This leads to two conclusions:

(1) the heat of shortening is not degraded work, since it changes sign when shortening becomes lengthening; and

(2) the viscous-elastic theory of the relation between force and speed must be altogether abandoned, since work is not absorbed, but degraded into heat, when a viscous system changes its form.

In fact, the process of shortening and doing work has the appearance of 'reversibility' in the thermodynamic sense. Irreversibility comes in only in the process of setting up and maintaining a state of activity. The maintenance of activity requires a continual flux of energy; but, apart from that, the process of shortening and doing work seems to have an exact counterpart in that of lengthening and absorbing work. When, however, stimulation is stopped the active state and all it implies rapidly disappear, a state of increasing plasticity sets in, and if the muscle is loaded it 'gives' and mechanical work is turned into heat.

The link between excitation and contraction

There is strong converging evidence that the process of excitation occurs at the surface of a muscle fibre, not throughout its substance. No known link exists at present between the events of excitation and those of contraction. This is curious in view of the relatively large amount of information about each separately. It has been suggested that excitation liberates certain chemical substances at the surface which then proceed to diffuse inwards, reacting with the contractile material on their way. The time, however, within which the active stage is found to be fully developed after a shock is far too short for such diffusion to have occurred. It is impossible, therefore, to suppose that the active state is propagated inwards from the surface by some actual substance diffusing in. A self-propagating process of some kind, started at the surface, must be looked for. The following statement by Dr B. Katz amplifies this.

(Communicated by B. Katz.) 'The capacity of skeletal muscles for speed and power depends not only on the intrinsic properties of the contractile matter, but also on a system of excitation by which the muscle fibre can be thrown into a sufficiently synchronous state of activity. It might be taken for granted that a maximum twitch is simply the result of a synchronous motor-nerve volley, yet the question remains how the nerve impulse arriving at a microscopic region of a muscle fibre can activate the whole of this fibre, say 0·1 mm. thick and a few centimetres long, with the required degree of promptness and uniformity. An important part of the answer is, of course, provided by the action potential of the muscle fibre which travels, from the end-plate towards the tendons, at a fairly high speed (approximately 2 m./sec. in the largest fibres of the frog at 20° C). In some long muscles of the frog this process has been shortened by the provision of two, or several, spaced nerve-endings on each fibre. But this impulse, like the nerve-axon spike, is a wave of surface activity, and while there is strong evidence that it does, in fact, evoke contraction in the interior of the fibre, the mechanism by which the two processes are linked is completely mysterious. It seems clear, nevertheless, that the contraction is a local sequel to the action potential, and that

it must start at the surface and quickly spread towards the centre of the fibre. The object of this note is to survey the facts on which this view is based.

'The twitch of a skeletal muscle fibre is invariably preceded by an action potential. If the action potential is blocked, for example, by local pressure or by application of an anaesthetic, then the contraction will travel up to the region of the block but no further. There is no evidence that the contraction as such can propagate itself along the fibre, but the action-potential wave is known to be self-relaying in nerve as well as muscle. The fact, therefore, that the twitch travels at the same speed as the action potential may be taken as evidence that the mechanical change is induced locally, at every point of the fibre, by the self-propagating electrical impulse.

'How are the two mechanisms coupled? The action potential gives rise to local currents which circulate through the muscle fibre. It might be reasonable to suppose that the movement of ions in the interior of the fibre in some way initiates the contraction. There is, however, a simple and cogent argument against this view. If an electric current is passed through a muscle from end to end, contraction does not occur at the same instant throughout, but an action potential starts at the cathode, and contraction follows it in its course. If the muscle has been paralyzed, by procaine, or by excess potassium or withdrawal of sodium, no propagation takes place, but a strong enough current still produces a local contraction in the neighbourhood of the cathode. It is clear that the flow of ions through the muscle substance is not in itself an adequate stimulus, but that contraction occurs at certain points, where the current leaves the fibre.

'Perhaps we might modify our hypothesis and say that a special component of the action current, namely, that which leaves the fibre, is capable of stimulating the contractile substance? There is, however, an alternative view, viz. that the stimulus arises not from the currents which circulate through the interior of the muscle fibre, but from the change of electric charge at the fibre surface. No final decision between these two possibilities is yet possible, but there are two indirect arguments which favour the latter. First it is doubtful whether an electric current alters the distribution of ions in a muscle fibre except at its surface membrane, and there is proof from experiments with internal microelectrodes that the critical ionic exchanges and the associated potential changes during the spike take place at the fibre surface and not, for example, at interfaces between fibrils and sarcoplasm. This makes it probable that the disturbance of the contractile matter must also originate at the fibre surface. Secondly, it is possibly to produce non-propagated contractions by chemically altering the membrane potential without any current penetrating the fibre at all. Certain types of chemical contractures, for example those caused by excess potassium or induced by veratrine, are associated with a depolarization of the fibre surface. Contractures of this type occur rapidly and are completely reversible, and they can sum locally with the effects of depolarizing (i.e. cathodic) currents, and be inhibited by repolarizing (i.e. anodic) currents. These reversible chemical contractures have an important property in common with the local cathodic contraction and with the normal propagated twitch, namely, that in all three mechanical activity is associated with a reduction in the electric charge of the fibre surface.

'It appears, therefore, that the requisite stimulus which initiates contraction is a diminution of electric charge density at the surface membrane. The question remains how electrical and mechanical events are coupled and how the stimulus is passed on to the contractile matter in the centre of the fibre.'

Volume changes

Changes of volume occur in muscle during and after contraction. They can be recorded accurately and quickly, and if a specific connexion could be established between them and chemical processes they ought to provide valuable information on the sequence of chemical events in contraction. Unfortunately, the muscle hitherto used, the gastrocnemius, has a structure mechanically so complex that during contraction a considerable pressure is developed inside it. The rapid reversible constriction of volume found to accompany contraction is, at any rate largely, the result of this pressure. Until volume change has been measured on muscles with long straight fibres it cannot be known whether the rapid reversible constriction has any significance; nor, if genuine volume changes do occur, what their time course and magnitude are. When a muscle has relaxed from a twitch a diminution of volume remains which presumably has accumulated during contraction; it is of the order of $5 \times 10^4\%$. During a tetanus, while the tension (and therefore, presumably, the pressure) remains constant the volume constriction goes on increasing. After relaxation the constriction usually diminishes (the volume increases again), corresponding, it is believed, to the resynthesis of phosphagen and the formation of lactic acid. In a muscle poisoned with iodo-acetate, in which lactic acid formation is prevented, the constriction may go on increasing after relaxation, corresponding to a continued breakdown of phosphagen.

In connexion with the present discussion of the contractile process (as distinguished from recovery) the only certain evidence available is that, after relaxation, there remains a constriction of volume which is of the order of size to be expected if the known heat production were derived from the breakdown of phosphagen or ATP; whether it is or not remains to be proved. The methods hitherto employed could certainly be improved and applied to muscles with long straight fibres. An obvious application would be to see whether volume changes could be recorded corresponding to the mechanical and thermal effects discussed above.

High pressure

High pressures (of the order of hundreds of atmospheres) can readily be applied to isolated muscle and the corresponding changes (mechanical, thermal, optical, electrical, etc.) observed and measured. The effect of a high pressure on the isometric twitch of a frog's muscle is considerably to augment and slow it. If the pressure is released at the moment when the shock is applied the augmentation disappears; if release occurs very shortly after the shock a large part of the augmentation remains. The effect of pressure on chemical reactions believed to occur in muscular contraction could be studied *in vitro*, particularly those in which volume changes occur; a comparison with the effects of pressure on contraction itself

might relate the two. The advantage of the technique is that pressures can be applied or altered very rapidly and do not injure the muscle. An obvious field may here await development.

Hydrogen-ion concentration

Most of the chemical processes involved in recovery cause, or are accompanied by, changes in pH. It would be very interesting to know whether the rapid processes accompanying a single twitch also involve changes of pH; this might give a clue to their nature. Unfortunately, no method hitherto used is capable of responding at the required speed. If pH changes occurring inside a fibre are to be recorded outside (e.g. by an electrode) they must be transmitted there by diffusion, and the speed of diffusion is so low that (quite apart from lag in the instruments themselves) the resolution of any rapid changes occurring in a single twitch would be extremely poor. It is scarcely possible to introduce a glass electrode into the inside of a muscle fibre in order to obviate diffusion—and even if one could do so, diffusion in the glass itself would probably introduce considerable delay, while the very high resistance might cause a long lag in recording. If a suitable indicator could be introduced, the colour change resulting from an alteration of pH, if large enough, might be recorded sufficiently quickly for useful comparison with the rapid mechanical changes involved in the transition from rest to activity, or with subsequent shortening. No present method, however, is nearly rapid enough, nor probably sensitive enough, to record changes of pH suspected to occur in a single contraction, except perhaps in a very slow muscle such as that of a tortoise at 0° C.

Electrical

The impedance of a muscle to alternating current of low frequency depends on the resistance and capacity of its surface membrane as well as on the resistivity of its interior. Changes of impedance can be observed when a muscle is stimulated, but unless the muscle is able to alter its external form they appear to be due only to changes in the membrane resulting from excitation. The membrane capacity is so large that to a current of higher frequency the resistivity of the muscle fluids is the only important factor, and there is no evidence of any significant change in this during contraction. The measurement of electrical conductivity is sensitive and rapid, and if significant amounts of potassium (for example) were freed or bound during contraction one could scarcely fail to detect the change. It would be worth while, therefore, to examine the matter again with an arrangement of multiple alternate electrodes, or otherwise, by which the effect of small changes of form, almost unavoidable in a contracting muscle, could be eliminated.

Osmotic

Osmotic changes undoubtedly occur and can be detected after a few contractions. The speed, however, with which they can be recorded is so low that little light is thrown on the process of contraction proper. The one important point which emerges is that the osmotic change is too great to allow us to suppose that the

precursors of the creatine, phosphate and lactic acid known to be produced during prolonged stimulation are themselves osmotically active. They must be present combined with something else.

Oxidation

The physical accompaniments of contraction and relaxation are unaffected by the complete absence of oxygen. It seems reasonably certain that oxidation is concerned solely with recovery.

Optical changes

Such scanty information as is available at present regarding the rapid sequence of changes which occur while contraction is developing or declining has been obtained by methods employing light in the visual range. The present speed of recording alterations of double refraction is too low to tell us anything more than that there is a reversible decrease of birefringence during shortening. Spectral absorption can be recorded with sufficient speed for the purpose, but characteristic spectra in the visible range are shown in measurable degree only by pigments connected with oxidation. No spectral change has been detected during a twitch, and the possibility of doing this is greatly reduced by the considerable alterations of opacity which occur. D. K. Hill (1949) has recently described the rapid changes of opacity which start in a twitch at the same moment as the earliest sign of the mechanical response. In a muscle under tension there is initially a sudden small increase followed rapidly by a decrease in light transmission, the decrease, and probably the increase, being due to alterations in scattering, not of absorption. Infra-red methods are being applied to the isolated constituents of muscle. If they could be made available for living muscle, and given sufficient speed, important information might be obtained as to the protein changes which occur during contraction and relaxation.

X-ray diffraction

The great difficulty in applying X-rays to study the rapid changes which take place in a single contraction is in the short exposure required—and, therefore, in the enormous power demanded for the X-ray tube. Possibly this could be avoided by slow repetitive stimulation with repeated short exposures timed to take place always at the same phase of contraction. In this way, for example, 300 twitches at 10 sec. intervals, each with an exposure of 0·1 sec., would give a total of 30 sec. exposure—if that is enough. If not, since a muscle will be altered by too much stimulation, the only way is to increase the power of the X-ray tube.

Thermodynamics

Following speakers will doubtless discuss the question whether contraction is to be regarded as due to entropy changes, as in the shortening of rubber, or on the other hand to changes of free energy due to chemical reactions. There are two pertinent facts to be faced: (1) If shortening is due to a muscle passing from a less probable to a more probable molecular state, i.e. from a state of low entropy to one of high entropy, as in the retraction of stretched rubber, the heat of shortening

should be negative. It is not, it is positive, namely (in mechanical units), about 400 g.cm. for every cm. of shortening of a muscle of 1 cm.² cross-section. If, therefore, we regard shortening as a thermodynamic process, since heat is given out in it the final entropy is less, not more. (2) When rubber shortens, if it is allowed to do work the heat absorbed is correspondingly greater. In a muscle, the heat of shortening a given distance is completely unaffected by work being done or not.

MUSCLE CONTRACTION AND MUSCLE PROTEINS

By H. H. WEBER

1. *Introduction*

There are three ways in which we can gain knowledge of the change of protein structure connected with muscular contraction. The first way is the observation of the contracting muscle itself. The changes of structural qualities and their time sequence have to be investigated. I take it that the excellent information which we have at present concerning the time sequence of mechanical, optical and energetic changes is one of the principal reasons for this meeting. Secondly to explain these phenomena we have to know also their localization. Up to the present we only have evidence that double refraction and probably also the changes of double refraction take place in the *A* band. The active change of length and tension in contraction may be also attributed to it.

Thirdly, we must know what really happens to the single particles of muscle proteins at the respective points and times. That is to say, we must know the properties and reactions of the purified proteins. Along this third way a large number of new facts have been established recently. I consider it to be my task to-day to report on the most important of these facts, emphasizing those which could throw light upon the structure and the properties of the living muscle. Consequently I will not speak on the 'corpuscular' proteins of muscle, the myogen- and globulin-*x* fractions.

2. *Components of myosin solutions and ATP*

The investigation of muscle proteins received an extraordinary impulse by the discovery of the interaction of myosin and adenosine triphosphate (ATP) by the groups of Engelhardt & Ljubimova (1939), Needham, Needham, Shen & Lawrence (1941) and Szent-Györgyi *et al.* (1942–8); the extracted and purified myosin fraction—I suggest calling it myosin briefly—splits ATP and is itself strongly influenced in all known properties by ATP. The enzymic effects of myosin might be due to the adsorption of enzyme proteins. This would seem to be almost certainly the case for the special ATP-ase studied by Polis & Meyerhof (1947). The effect of ATP on the colloid properties of myosin, however, is a basic process; the formation and dissociation of actomyosin out of two single proteins. Actomyosin consists in its greater part of that protein which Szent-Györgyi now calls myosin, and in

a minor part of Straub's actin. I shall henceforth call this myosin *L*-myosin, according to the first discoverers (Schramm & Weber 1942), in order to avoid confusion with the total fraction and its varieties (myosin *A*, myosin *B*).

Figure 1 shows the sedimentation constants of the various purified myosins as a function of protein concentration. Curve 1 gives pure undenatured *L*-myosin, curve 1*a* denatured *L*-myosin, originating *directly* from *L*-myosin without any
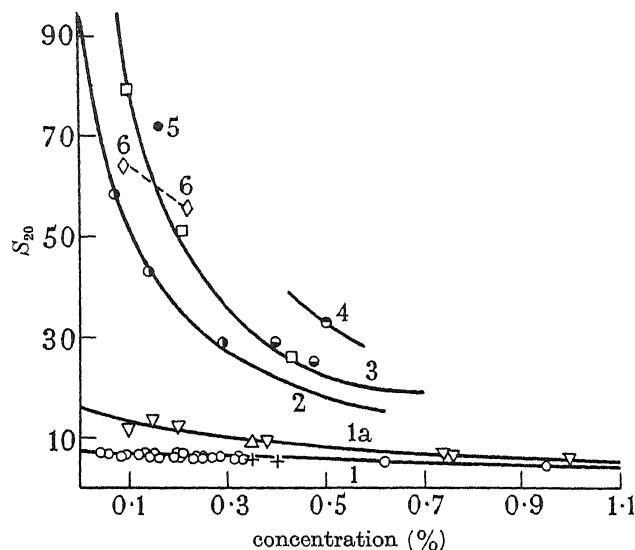
FIGURE 1. Sedimentation constant as a function of concentration. Curve 1: *L*-myosin ○ purified by fractionated precipitation, + by crystallization, ◒ ◖ by ATP dissociation of the actomyosin of curves 2 and 4. Curve 1*a*: denaturated *L*-myosin △ homogeneously, ▽ 50 % mixed with undenaturated *L*-myosin. Curves 2, 3 and 4: actomyosins ◒ □ ◖ isolated by fractional precipitation from the muscle extract. ● 5 composed of *L*-myosin of curve 1 and actin. ◓ reassociation after the splitting of ATP. Curve 6: isolated *F*-actin.

intermediate stages. Both proteins have the same viscosity and both alike are indifferent to ATP. The circles of curve 1 represent *L*-myosin purified according to Weber, the crosses indicate crystallized myosin according to Szent-Györgyi. 2, 3 and 4 are actomyosins separated from *B*-myosin extracts. The points in curve 1 resembling those of 2 and 4 give the sedimentation constants of these actomyosins after addition of ATP. The ◓ points of curve 3 indicate sedimentation constants of preparation 4 after spontaneous splitting off of ATP by the myosin. Point 5 marks the sedimentation constant of an actomyosin prepared by mixing pure *L*-myosin and *F*-actin. Points 6 are sedimentation constants of some preparations of *F*-actin. These curves first confirm the important conclusions arrived at by the Szent-Györgyi group (1942–8) with a method which allowed the homogeneity of the preparations to be judged. Actomyosin consists of pure actin and *L*-myosin and is reversibly split into both components by ATP. Furthermore, the curves show the existence not only of one actomyosin but of several. Their formation apparently takes place in sharply separated morphological stages. This is very clearly shown by the fact that, before the isolation of the single actomyosin, one usually gets preparations which contain two or three of these actomyosins side by

side (for instance, in the researches of Snellmann & Tenow (1948) with B-extracts). This last fact disproves the viscosimetrically founded assumption of the Szent-Györgyi group, that mixtures in any proportion combine to a uniform actomyosin of different composition and size of the particles.

The increase of the sedimentation constants of L-myosin through the different actomyosins indicates an increase in the thickness of the particles, because the lengthening of rods alone does not greatly affect the sedimentation. Thus the thickness of the several actomyosins differs, and all are much thicker than L-myosin.

That actomyosin particles also differ in length and are always longer than L-myosin particles is already clear from the viscosity studies of the Needham and Szent-Györgyi groups. Figure 2 shows the curves of viscosity relative to the velocity gradient, the preparations being the same as in figure 1.
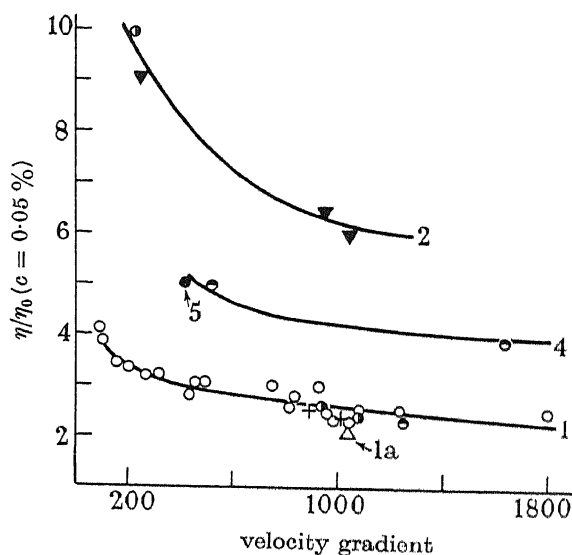
FIGURE 2. Viscosity as a function of velocity gradient. Description of preparation as for figure 1.

It is remarkable, however, that the viscosity of actomyosin, curve 4, is lower than that of the preparation of curve 2, although the sedimentation constant is higher. The particles of preparation 4 are shorter and thicker, those of preparation 2 longer and thinner, but both are fairly monodisperse. These differences await explanation. The variability of the shape of the actomyosin particles must be kept in mind because the Szent-Györgyi group has estimated the actin rate of muscle based on the viscosity of their extracts. Therefore it might be a good thing to control this estimate in another way.

As theory progresses the differences in light scattering between actomyosin and L-myosin may possible become highly important. They are considerable. The apparent extinction of actomyosin solution of 1 % equals 1 cm.$^{-1}$; of L-myosin of 1 % is 0·05 cm.$^{-1}$. We intend to develop these studies systematically.

We know little of how the single actin and myosin particles associate. Bailey & Perry (1947) have pointed out that the sulphydryl groups are of importance.

But we do not know whether the myosin and actin particles in forming actomyosin associate side by side, though this is assumed by Szent-Györgyi (1943). But actomyosin threads of muscle extracts in electron micrographs (Hall, Jakus & Schmitt 1946) are not thicker than threads of pure *F*-actin. Aggregation lengthways would suggest that the long *F*-actin threads dissociate intermediately into *G*-actin particles. But in actomyosin dissociation by ATP in solution, it is *F*-actin, not *G*-actin, that is always set free at once (Kaumanns, unpublished).

Even in extremely low concentration, actomyosin coupled with fluorescent dyes is able to produce phosphorescence, which decreases in dissociation. Several theoretical explanations are possible.

3. *Shape and size of the* L-*myosin particle*

While knowing rather little of the beginning of the process leading from actomyosin to actin and *L*-myosin we know somewhat more of its end; Dr Hilde Portzehl (in the Press) has completed her measurements of sedimentation (figure 1)
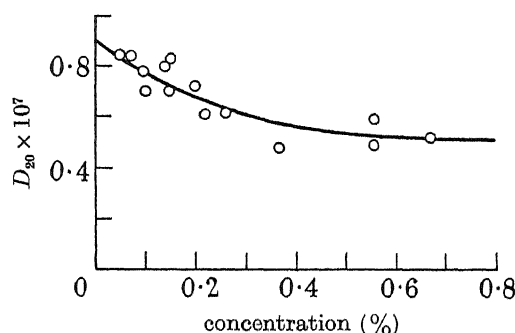
FIGURE 3. Diffusion constant of *L*-myosin as a function of concentration.
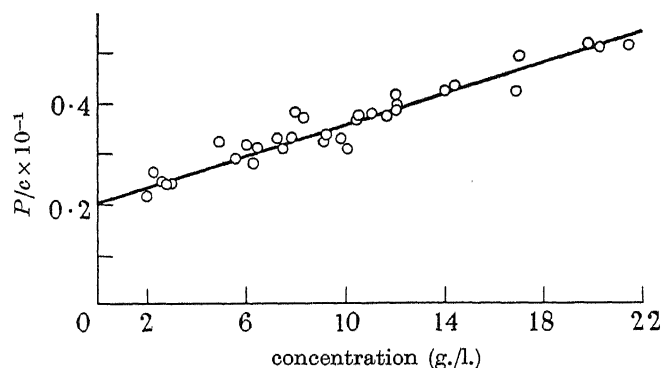
FIGURE 4. Osmotic pressure measurements on *L*-myosin.

by measurements of diffusion. From the values of $S_{20}(c = 0)$ of 7·1 and $D_{20}(c = 0)$ of 0·9 there results a particle weight of 821,000 and an axial ratio of 94.

She has further measured the osmotic-pressure protein concentration dependence. The values of the measurements lie between 0·04 and 1·02 mm.Hg. The particle weight calculated from the extrapolated value of P/c is 840,000 ($\pm$ 33,000)

and from the formula of Schultz (1947) $q = \dfrac{\rho B}{A \cdot 0 \cdot 785} \times 10^{3*}$, an axial ratio of 128 is calculated.

From the mean axial ratio and the weight of the particles the particle length is 2300 Å and the thickness 23 Å, assuming a specific volume of L-myosin of $0 \cdot 74$.
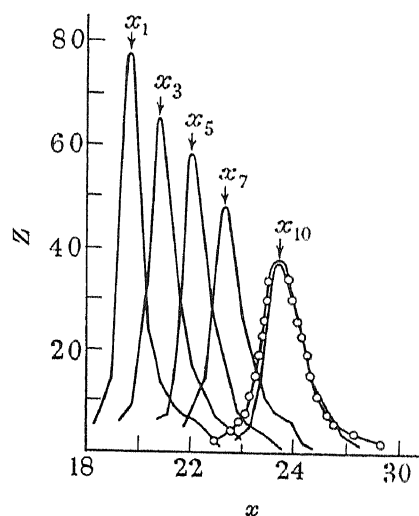
FIGURE 5. Sedimentation gradients of L-myosin. —— sedimentation gradients observed; O—O x_{10A} calculated from curve x, and $D_{20}(c = 0 \cdot 25 \%)$. $Z = $ scale-line displacement; $x = $ distance from centre of rotation.
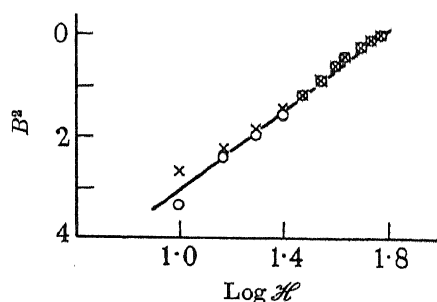
FIGURE 6. Symmetry of the diffusion gradient in L-myosin. $B = $ width of the gradient curve; $\mathscr{H} = $ height corresponding with width measured.

Figure 5 shows the observed curves of the sedimentation gradient $x_1 - x_{10}$. Curve x_{10A} is found by calculating how much curve 10 must be wider by diffusion than curve 1. On the left side 10_A is even somewhat wider than curve 10 due to the dependence of the sedimentation on the concentration. Thus the sedimentation is strongly monodisperse.

The diffusion of L-myosin also is monodisperse; for the square of the width of the gradient curve depends on the logarithm of the height at which the width is measured (figure 6).

$$* \quad q = \frac{\text{length}}{\text{thickness}}; \quad \rho = \text{density of protein}; \quad A = \frac{p}{c_{(t=0)}}; \quad B = \frac{\Delta p/c}{\Delta c}.$$

A particle weight of 70,000 and a moderately asymmetric shape would be plausible for G-actin with a sedimentation constant 4 and a smallest tryptophan unit 70,000.

F-actin seems to consist of threads of a considerable and varying length of from 1 to several μ according to the electron microscope pictures made by Jakus & Hall (1947). That, nevertheless, the sedimentation is uniform could be either due to their constant diameter (§ 2) or to the fact that F-actin particles in solution are smaller and more uniform than in their pictures. For the process of high vacuum drying and temperatures of several hundred degrees in an electron microscopic object are not without danger.

Table 1

protein	particle weight			axial ratio*			dispersity	length (Å)	width (Å)
	ultra-centrifuge	osmotic pressure	mean	ultra-centri-fuge	osmotic pressure	mean			
tropomyosin (Bailey 1948)	0.92×10^5	0.88×10^5	0.9×10^5	56	111	84	homio	990	11.8
L-myosine (Portzehl in the Press)	8.2×10^5	8.4×10^5	8.3×10^5	94	128	100	homo	2300	23
G-actin	—	—	$\sim 0.7 \times 10^5$	—	—	~ 5	homio	~ 14	~ 2.8
F-actin (Jakus & Hall (1947)	—	—	—	electron microscope			poly	$\geqslant 10^3$	80–140
actomyosin (Schmitt et al. 1946)	—	—	—	electron microscope			poly	$\geqslant 10^3$	50–300

* Uncorrected for hydration.

4. Gels and threads of actomyosin

The solutions of myosin proteins set to a gel if the ionic strength decreases sufficiently. Only pure actin always remains in solution changing from F-actin to G-actin (Straub 1943). At pH 6·7 this occurs in actomyosin very sharply at 0·28 to 0·27 μ (Portzehl, Schramm & Weber 1943). The less sharp change of L-myosin is finished at 0·05 μ. If F-actin and L-myosin are separated in solution under the influence of ATP they will combine as soon as the ionic strength falls below 0·2 (Szent-Györgyi 1947). For a gel is formed at an ionic strength too high for pure L-myosin and the supernatant liquid no longer contains free actin. Such a gel containing ATP shrinks to a protein concentration of 50 % and more if the ionic strength continues to decrease to values below 0·1 to 0·15 μ (Szent-Györgyi 1943). The same shrinking occurs if ATP is afterwards added to an actomyosin gel of this ionic strength (0·05) free from ATP (Szent-Györgyi 1942). If the gels are somewhat orientated by tension—stretched myosin threads—they contract instead of shrinking (Szent-Györgyi 1942; Buchthal 1947).

The shrunk gel swells again to the initial volume if the ionic strength is again increased to the initial amount—but only if ATP is added. The equilibria seem to depend reversibly on the ionic strength. ATP, however, seems to be necessary so that these equilibria are really obtained. Under the influence of ATP, myosin

threads become extraordinarily extensible (Buchthal, Deutsch & Peterson 1947). Already in the case of small tensions they lengthen under conditions under which they would otherwise shorten. The changes of double refraction may find a satisfactory explanation if we assume the threads to be incompletely orientated.

All the experiences of the effects of ATP on defined myosin proteins can be summed up as follows: ATP diminishes the forces of cohesion between these proteins. In solution this diminution leads to complete dissociation. In a gel this cohesion is diminished so much that the extensibility rises enormously and the resistance to shrinking, contraction and swelling, which may themselves be caused by the ionic strength, decreases.

5. *Specificity of the reaction between ATP and myosin*

If this result seems to be somewhat meagre, we must consider that the complex of these phenomena is so extraordinarily specific that it is difficult not to assume a connexion with the phenomena of muscle activity; for only ATP sets free all these phenomena. And vice versa, only the contractile actomyosin is influenced in the manner described. The other muscle proteins especially are not sensitive to ATP, apparently not in the muscle itself either (see Kamp (1941) and compare also Dubuisson (1948) and Szent-Györgyi *et al.* (1942–8)). The Meyerhof reaction between proteins and lactic acid lacks this high specificity.

A diminution of the cohesive forces by ATP at the moment of contraction would bear out the idea that muscular contraction is due to the formation of a thermokinetic skein of molecular chains (Woehlisch 1940). In this the decrease of the cohesive forces would increase the thermokinetic mobility during contraction. It is difficult to make this conception harmonize with many experimental results.

6. *Contraction of water-extracted fibres*

So far the analysis gives no convincing explanation of the vital contraction. Indeed, the difference between the contraction of the myosin thread and that of the muscle does not even enable us to decide whether ATP acts on actomyosin at the moment of contraction or afterwards.

This uncertainty disappears entirely or almost entirely if we regard the next model of Szent-Györgyi (Varga 1946), namely the water-extracted, frozen, and long-stored muscle. This also contracts under the influence of ATP even in solutions of K salts of ionic strength 0·4 to 0·5. This shortening reaches its maximum at temperatures of 10 to 12° C; it diminishes at lower temperatures and ceases completely at 0° C. The maximum shortening at physiological ionic strength is 60 to 70 % of the initial length. Isometrically a tension of 1 to 2 kg./cm.^{2} is developed (Szent-Györgyi 1949). According to a letter which I received from Meyerhof this fact has been repeatedly confirmed, for instance, in the laboratory of D. Nachmannsohn. Nor is this mere contraction in length reversible if the fibre bundle is put into KCl solutions up to 2 μ (own researches). Here ATP itself seems to act as a shortening substance.

The mechanism is not yet known. Szent-Györgyi has tried to calculate the change of free and total energy from the dependence on temperature. In my opinion,

however, it has first to be established that a true equilibrium capable of attainment from both sides solely by a change of temperature, is being studied.

It is not yet known why ATP produces a model contraction in the washed-muscle system but not in the actomyosin thread. No doubt the orientation of the particles in the contractile myosin thread is greatly inferior. Moreover, there is no doubt that the muscle model also contains, in addition to actomyosin, all the other proteins insoluble at low ionic strength. As such we may probably count tropomyosin, so excellently characterized by Bailey (1948) as well as the N-protein of the Szent-Györgyi group and also the insoluble stroma proteins.

In order to decide whether the decisive factor is the different orientation or the different composition of the structure of the two models, it would be well to know how the above-mentioned proteins combine to compose the insoluble muscle structure. A bridge can be thrown between the living muscle and the washed-muscle model, or the system of pure actomyosin, only when the composition of the insoluble muscle structure is better known. We have started corresponding researches.

7. Comparison of the phenomena of muscle and contracting models

The bridge from the reported results to the vital contraction is narrow so far as useful facts are concerned, but it is broader with regard to the possibilities of future research. The fact that the ATP effect seems to be connected with a diminution of the cohesive forces might agree with the suggestion of Sandow (1947), according to which the latency relaxation is due to a reaction of the contractile structure with ATP.

But it would be a result of extraordinary importance if we were entitled to regard the ATP contraction of the washed muscle as a contraction without relaxation. The correctness of such a view would be the more guaranteed, the better the other colloidal phenomena of the vital contraction agree with the corresponding phenomena of the models.

A consideration of the elastic changes does not reveal much because we do not sufficiently know the elasticity of completely oriented contractile models.

Changes of double refraction are found in vital contraction as well as under the influence of ATP on all models and systems of actomyosin. But this phenomenon is of little significance as long as we do not know to what degree the changes of double refraction have to be attributed to the rod or to the intrinsic double refraction.

The muscle itself has not been found to show changes of phosphorescence (Szent-Györgyi et al. 1947).

Also the present results of light scattering and its changes offer several explanations. The magnitude and the change of the total intensity of scattering depend on two facts: the scattered intensity becomes smaller if the scattering particles become smaller by dissociation, but also if the orientation of the particles become more regular. In solutions these two influences can be separated by the investigation of concentration series, but not in muscle. Only in rare cases will an indirect decision be possible. A disorientation might come into question, when, for instance,

v. Muralt (1935) points out that the muscle on the formation of lactic acid scatters more only if it is allowed to swell and that, on the other hand, every swelling leads to increased turbidity.

The scattering pictures become much richer and more characteristic, if besides the intensity of the total scattering their dependence on the angles is also measured. In this case, as the excellent work of Buchthal & Knappeis (1940) shows (compare also Schaefer & Göpfert 1937) the diffraction by microscopic structures and the scattering by colloidal particles overlap. That also the colloidal elementary particles contribute to this dependence on the angles must be expected from the mere fact that at least one diameter of the protein rods should always be larger than $\frac{1}{10}\lambda$, for instance in L-myosin $\frac{1}{2}\lambda$. A distinction of the different influences and interpretation seem not to be impossible, especially if by way of comparison also the corresponding aberration of light in the different systems of actomyosin are investigated and taken into consideration.

The other phenomena accompanying the beginning of muscle activity are so well established in the muscle itself, that it would be more useful first to examine them adequately in the actomyosin gel and especially in the washed muscle. Even the washed muscle with its unknown microscopic structure offers two advantages: first, it has no metabolism. Any alteration taking place on adding ATP has its cause in the addition of ATP; that is especially the case for the performance of work and the change of the mechanical and elastic properties. Since the contraction of washed muscle is in the nature of a contracture, methods which cannot be applied to a single twitch can here be used, e.g. the analysis of the intrinsic and rod double refraction and the X-ray diffraction.

The co-ordination of Szent-Györgyi's muscle model and the pure actomyosin systems may finally lead to the knowledge of what happens to the single particles of the contractile protein.

X-RAY STUDIES OF MUSCLE

By W. T. ASTBURY, F.R.S.

No X-ray diffraction photographs have yet been obtained of muscle in the very act of contracting or relaxing, but much work has been carried out in the approaches to this goal. I reviewed the situation in 1945 in my Croonian Lecture (Astbury 1947 a), and the present contribution may serve partly to recall some of the main conclusions already arrived at then, but chiefly to develop them in relation to more recent discoveries.

In all the kinds of muscles that have so far been examined by X-rays the dominating feature has been found to be the fibrous protein myosin lying along the direction of the myofibrils; further structural components have been revealed according to the type of muscle, it is true, but myosin at least seems to be essential, and there are very strong reasons, not only from X-rays but from much other evidence besides, for believing that it is indeed the mainspring of the contractile apparatus proper. Many X-ray and accessory studies have been carried out on it

both *in situ* and after extraction, and it is a principal inference from such investigations that, whatever the detailed steps by which contraction and recovery are contrived in muscle itself, myosin has at any rate properties peculiarly fitted for these purposes. The myosin molecule or complex is inherently elastic; it can be caused to change its internal configuration and thereby alter its length over a large range.

X-ray studies have brought to light something of much wider significance than this, however, for it is now seen that the basic plan and properties of the myosin complex are by no means unique; and by inference, muscular activity, when reduced to its simplest terms, also appears as probably no more than a special case of a very fundamental scheme that has been adapted to many uses. Myosin is only one member of a *family* of fibrous proteins which includes also keratin, epidermin, fibrinogen and fibrin, paramyosin and tropomyosin, and even the structural protein of certain bacterial flagella. All these are characterized by a common type of large-angle X-ray diffraction diagram, and they share, too, a number of other properties that are correlated with the X-ray manifestations. In their normal state they are constructed of polypeptide chains in a regularly folded configuration—what we have called the α-form; but they can be pulled out from this shape into the extended configuration that we have called the β-form. As is now well known, this reversible intramolecular transformation is the basis of the long-range elasticity of mammalian hair—or, for that matter, of oriented strips of extracted myosin. In addition to their capacity for extension, and of still greater moment, the members of the k-m-e-f group have a remarkable power, under suitable treatments, of contracting well *below* their normal α-length—the property that we have called supercontraction. From the viewpoint of X-ray analysis and supporting studies, supercontraction is the master-key to the interpretation of muscular contraction. We infer that just as members of the k-m-e-f group can on the one hand be extended from their normal folded α-configuration, so on the other can their constituent chains be caused to shorten by folding even more compactly than in the α-form. The whole process of contraction and elongation is thus a question of states of folding of polypeptide chains. There is little doubt now that preparations of extracted myosin can be brought into various states of folding in the laboratory, but the immensely difficult problem we face is what are the exact physical and chemical steps involved in a muscle itself. This is nothing less than the problem of protein structure in general, and it is hardly likely that an answer of any completeness will be forthcoming in the foreseeable future.

What seems to be the most important recent addition to our knowledge of the partners of myosin in the muscle cycle is Straub's (1942, 1943) discovery of the new protein actin. From this work and that of Szent-Györgyi (1942, 1947) and his collaborators it has come to be accepted that, strictly speaking, it is not myosin alone but the complex, actomyosin, that it forms with actin that is the structural substance of myofibrils and the seat of their contractile properties—in skeletal muscle at least. Extracted actin is capable of assuming either of two modifications: it can exist either in a corpuscular (or globular) form, *G*-actin, or in a fibrous form, *F*-actin, and the transition from the former to the latter may be brought about,

reversibly, by lowering the pH or by adding KCl. Examination in the electron microscope, by Jakus & Hall (1947) and by Astbury, Perry, Reed & Spark (1947), shows that F-actin consists of long, fine fibrils built up by the linear aggregation of the corpuscular units of G-actin; and we have also succeeded in obtaining X-ray fibre photographs of F-actin (Astbury *et al.* 1947; Astbury 1947*b*) from which the period along the fibrils is found to be 54 Å or a multiple thereof—probably 108 Å. These X-ray diagrams of actin at once resolve a long-standing problem in the interpretation of the diagram given by the frog sartorius muscle, for example. The complete diffraction pattern of such a muscle (unfortunately not in the living active state, but after being allowed to dry) is now seen to be effectively the sum of the patterns given by extracted myosin and F-actin separately; and the two are oriented parallel to each other and to the muscle-fibre axis (Astbury 1947*b*, 1949).

We have here then X-ray 'visual' evidence of the two main components of what is supposed to be, and very probably is, the actual centrepiece of the contractile apparatus—again in skeletal muscle at least. We have no information yet in what state actin exists in living muscle, and especially during the times of contraction and relaxation, but present biochemical indications from such investigations as those of the Szent-Györgyi school and of Bailey & Perry (1947) are that it goes in and out of combination with myosin through the intervention of ATP. Exactly how and where the stages of forming and loosening the actomyosin complex in this way fit into the contractile cycle is still not clear, but the chief point I should like to make in this connexion is that there is nothing in these findings to invalidate the X-ray interpretation I have just summarized; rather do they supplement it by filling a gap in the argument and supplying just such a chemical mechanism as is required to bring about the inferred configurational and length changes in the myosin component. I should like to suggest that it is when the myosin and actin *combine* that the myofibrils shorten; and the reason for this is the oldest but still the most impressive fact about muscular activity—that on stimulation the structure passes over suddenly from a semi-fluid and plastic-like consistency to a state almost like that of a crystalline solid. In the traditional phrase, it becomes 'like steel'. The general idea will be apparent from figure 7, in which, in the relaxed state, the myosin chains and the actin corpuscles are depicted as relatively free, the former being perhaps partly in the β- as well as in the α-configuration; while in the contracted state the system is supposed to aggregate by the side-to-side union of linear arrays of actin corpuscles, i.e. F-actin, with myosin chains in a supercontracted configuration. A picture of this kind would appear not only to harmonize existing X-ray and biochemical knowledge but also to be in accord with certain thermodynamic tests which indicate that whereas the entropy factor plays a controlling part in the elastic properties of relaxed living muscle, the internal energy factor predominates at the moment of stimulation and contraction.

Important advances have been made recently in studies of striated muscle with the electron microscope. An outstanding conclusion arrived at by Hall *et al.* (1946; Schmitt *et al.* 1947) was that the finest resolvable myofibrils run straight on through all the bands and still remain taut after contraction. We had already inferred from X-ray studies on living plain muscle that the myosin chains preserve in the main

their orientation over considerable changes of the muscle length, and this new electron microscope evidence, with its verdict against any visible coiling or crumpling, confirms that we must look for configurational changes at the level of the molecules themselves. The series of bands comprised within each sarcomere of a striped muscle is beautifully shown up in electron micrographs, particularly those of Draper & Hodge (1949), who have also succeeded in revealing the presence, throughout the whole length of the sarcomere, of additional, very fine striations
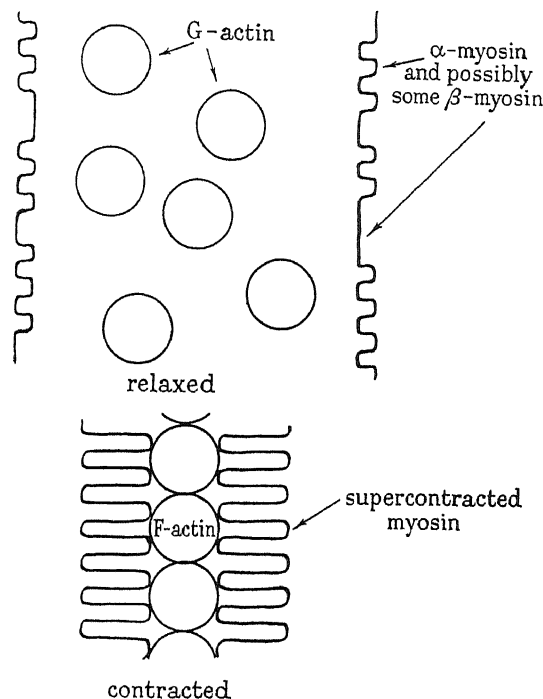
FIGURE 7. Diagrammatic representation of a possible mechanism of muscular contraction by the formation of an actomyosin complex.

spaced at about 400 Å. The same workers have developed, too, a micro-incineration technique using a high intensity of the electron beam itself, and by this means they have shown that, while the Z and M bands disappear on incineration, certain other features and especially the fine striations remain. They conclude, among other things, that the latter are very probably associated with a periodic concentration of magnesium and calcium bound to the continuous actomyosin framework. There is here an attractive link with the X-ray data on actin, for on the one hand Straub has shown that the actin molecule includes calcium in its make-up, while on the other Bear inferred some years ago that the small-angle reflexions in the X-ray diagram of frog sartorius (which we realize now, as mentioned above, to arise in some way from the presence of a regular system of actin corpuscles) possibly require a full fibre period of 350 to 420 Å (Bear 1945; Schmitt *et al.* 1947).

Reverting once more to the place of muscular activity in the wider and deeper cytological relationships of the k-m-e-f group as a whole, I should like to add that the concept of the essential muscle machine as a combination of myosin and actin

suggests a remarkable adaptive economy in the uses to which the members of this family of fibrous proteins has been put. The key facts are these: (a) the epidermis, itself composed largely of fibrous protein of the α-type, gives rise to two kinds of keratin: mammalian α-keratin and the so-called 'feather keratin' characteristic of birds and reptiles; (b) α-keratin is a molecular analogue of myosin; and (c) the X-ray fibre diagram of feather keratin is both an orthodox fibre diagram and also what may be called a corpuscular fibre diagram of the type of F-actin. The implication, therefore, is that the dual function of the epidermis, whereby it produces structures sometimes of the α-type and sometimes of the feather type, is unified in muscle to build a 'two-component machine'. The complete X-ray picture of hair keratin, as exemplified for instance by MacArthur's (1943) diagram of porcupine quill tip, tells the same story from another aspect. The correspondence between the full diffraction patterns of porcupine quill and frog sartorius is so striking that there can be hardly any doubt that fundamentally the same molecular plan underlies them both. The muscle diagram is a composite pattern arising from myosin plus actin; the porcupine quill pattern would appear to arise similarly from α-keratin plus a keratinous analogue of actin. It is as though there still persists in hair and such-like the framework or debris of a muscle machine, but the switches are lost and the fire has gone out, and the apparatus can be made to work now only crudely by pulling it about and by other imperfect laboratory devices.

The very latest addition to the hair and muscle family is the fibrous protein of certain bacterial flagella—actually *Proteus vulgaris* and *B. subtilis*—which we have been examining by X-rays in collaboration with Dr C. Weibull of the University of Uppsala (Astbury & Weibull 1949). The investigation is still proceeding, but already it is clear that these flagella, which behave physico-chemically like long macromolecules about 120 Å thick, fall into the same X-ray group as keratin, myosin and the rest, and that from this viewpoint, at least, there is considerable justification for thinking of them as of the nature of *monomolecular* hairs or muscles. This is equivalent to saying that they might correspond to the contractile *units* of muscle tissue, though of course not necessarily activated in the same way. It is a principal deduction from the X-ray and related findings that just as we must not look upon the muscle protein myosin as unique in plan, so we ought not to consider the chemical cycle by which myosin is activated in muscle as anything more than an adaptation to a special purpose. Here in the bacterial flagella, with their intense motility, we have another of the various analogues of myosin, but presumably brought into action by different means. It is not an easy approach to the muscle problem, this of the movements of bacterial flagella, to be sure, but at any rate it has all the signs of bringing us one step nearer the core of the matter. Unless the flagella are whipped simply from their roots, the sort of picture that comes to mind is of a contractile phase passing rapidly down one side and up the other—as if, in fact, there were a long thin polypeptide loop, of the type of keratin and myosin, along which a change in the state of intramolecular folding, once initiated, ran back and forth as a wave.

One obvious experimental lesson emerges from these X-ray results thus briefly summarized, and that is that it is imperative to be able to carry out a similar

detailed exploration of the diffraction diagram of muscle in the living state, and especially at the actual moments of contraction and relaxation. It is not that we have been unable at all to take X-ray photographs of living muscle; we have succeeded in doing this for steady contracted states of plain muscle and the experiments are described in my Croonian Lecture, but it is the merest beginning of the task. In the first place the permitted exposure was not long enough to reveal much more than the outlines of the myosin component—the actin contribution is fainter and in any case requires long-distance high-resolution photographs—while in the second, the muscle was not actively contracting or relaxing. An estimate of the difficulty may be formed from some of the exposure times taken by MacArthur, even with our high-power rotating-anode tube, in his studies of porcupine quill tip —a dry, dense, and unusually crystalline specimen of keratin, be it noted; these exposure times often amounted to the order of 20 hr. Suppose now, as suggested by Professor Hill, we were to stimulate a muscle every 2 sec. for 5 min. and isolate by means of a shutter synchronized with the muscle a phase of 0·05 sec. duration in each twitch, the total exposure would be still only 7½ sec. Nevertheless, in spite of such discouraging calculations based on present X-ray equipment, I firmly believe the necessary experiments could be carried out and should be tackled. It would involve building a super X-ray tube that would deliver a very intense 'flash' beam triggered off by the contracting muscle itself, and it would require physicists and biologists working in close co-operation. At the moment there is no apparent prospect of performing the experiments without special support and co-ordination, but it is a worthy project without which further really substantial X-ray progress into the details of the muscle machine is hardly possible.

Some chemical and physical aspects of muscle contraction and relaxation

By M. Dubuisson

We have been interested in three questions in relation to the problem which is under discussion to-day: the impedance changes, the pH changes and the protein changes during muscle contraction and relaxation.

(a) *Impedance changes*

With high-frequency current, we were not able to find any change in the conductivity of the muscle cell (Dubuisson 1935, 1936, 1937 a). This confirms the earlier discoveries of MacClendon (1929) and Hartree (1933). At low frequency, changes were found due to modifications of some interfaces in the muscle cell. There is first of all a very fast decrease of impedance (*a* wave) which occurs before the contraction and which is synchronous with the action potential (figure 8). Some years later, K. S. Cole (1938) found the same phenomenon in nerve.

During contraction and relaxation, two increases of impedance were found, the first (b_1 wave) starting with the shortening, the second (b_2 wave) with the relaxation (figure 9).

No direct relations between these phenomena and the chemical events could be found, except that when the muscle is poisoned with monoiodoacetic acid (so that no lactic acid is formed and practically no phosphagen resynthesized) the increase
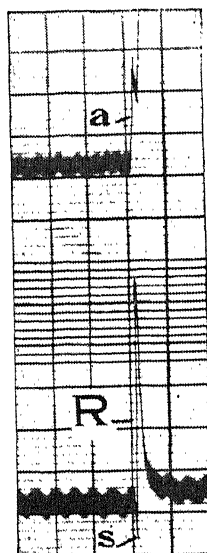
FIGURE 8. *Above*: decrease of impedance (*a* wave) preceding the mechanical contraction of a frog's sartorius. *Below*: action potential (*R* wave) recorded simultaneously.
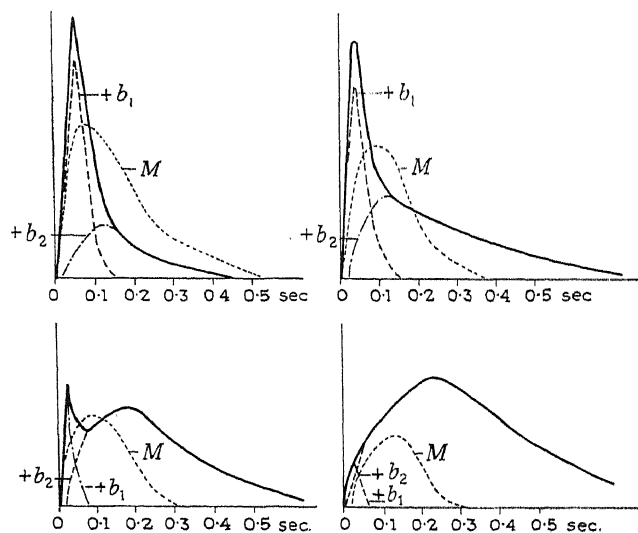
FIGURE 9. Schema of four records during fatigue. Full line: impedance changes during the contraction of a frog's sartorius; dotted M, myogram, taken separately and drawn on the same scale of time. In $+b_1$ and $+b_2$, supposed curves of the two impedance changes; temp. 18° C, direct stimulation.

of impedance which occurs in relaxation becomes an irreversible phenomenon. This is probably due to the fact that the physical properties of the interfaces responsible for the last wave are in close relations with the metabolism of phos-

phagen, but we do not know anything about the nature of those interfaces nor their location.

There were some reasons to believe that the pH changes due to the metabolism of some substances involved in contraction and relaxation could be responsible for some of our impedance change pictures. This was the starting point of our work on pH changes during muscle contraction.

(b) pH *changes*

. With a special glass electrode and a very sensitive electrometer, we succeeded in getting records of pH changes (Dubuisson 1937 b, c, d, 1938 a, b, 1939 a, b, c, 1940). It took a long time before we could be quit of muscle potentials in our records, but we succeeded in doing it. The principle of the technique is based upon the fact that if a muscle is in equilibrium with a solution which contains a certain amount of bicarbonate and CO_2, if we take such a muscle out of the solution, a thin liquid film is left on the surface of the muscle. If we now press a glass electrode of a special form on this film, we can measure the pH of it. Now, if any change of pH occurs inside of the muscle fibre, owing to the presence of bicarbonate, the tension of CO_2 will be decreased or increased, not only inside the fibre but also in the interfibre spaces and in the film of solution lying under the electrode. There are, of course, quantitative relations between those pH changes and the amount of CO_2 going out or coming in, and calculations may be done if we know the retention factor of the muscle cell. The only trouble is that the diffusion of CO_2 is not instantaneous, and as a result of this the records are delayed, at least when the changes are fast. Therefore, we began our work with the smooth muscle of the stomach of the Hungarian frog in which the latent period is at least 5 sec.

More than one hundred records were taken in different conditions, including monoiodoacetic poisoned muscles. We may conclude from all those records that there are four changes of pH. The first two started almost at the same time (a alkaline and b acid), 2 to 8 sec. after the beginning of stimulation and 2 to 4 sec. before any contraction can be seen (figure 10). These two first changes must have occurred still earlier inside the muscle fibres, so that there is no doubt that there exists an alkaline and an acid change which *both start during the latent period of the tissue*. The acid change is much more important than the alkaline one, greater when the tension is high, smaller in fatigued muscle. During relaxation, there is another alkaline change (c wave) and still later an acid one (d wave). The two last changes are due, the alkaline one to the splitting of phosphagen, the acid one to the formation of lactic acid; pH records and chemical analyses made in Meyerhof's laboratory in Heidelberg in 1937 were indeed in very good agreement. The maximum of the pH increase due to phosphagen splitting comes, on our records, at the time of relaxation. The maximum of the pH decrease due to the formation of lactic acid is only reached 1 or 2 min. after the muscle has returned to rest.

There exists a delay between the pH we record and the real pH inside muscle fibres. The error of time is relatively greater when the changes are fast. This means that the two first changes *come earlier inside the fibre*, and that there consequently

exists an alkaline phenomenon and an acid phenomenon preceding the mechanical shortening of the smooth muscle.

The fact that the precontractile acid phase is unaffected by poisoning with monoiodoacetic acid, that it comes before the alkalization due to the splitting of phosphocreatine, that this acid phase is much more important when the initial pH of the muscle is higher, cannot be of course a proof that it is due to the splitting of
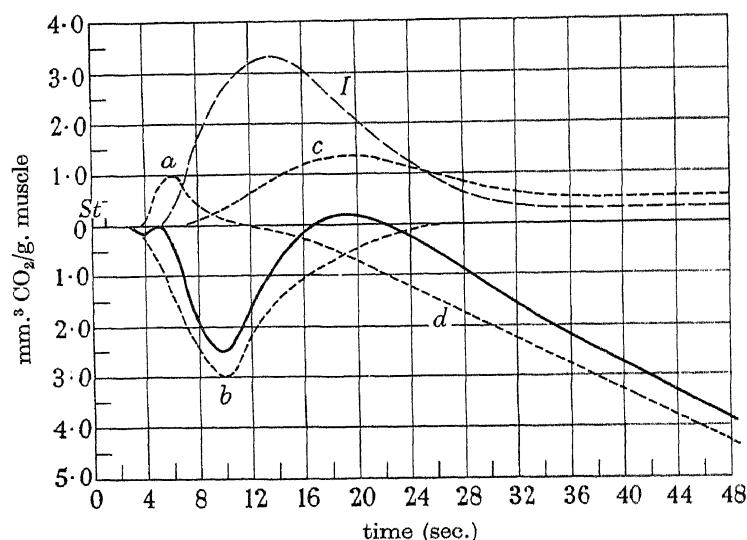
FIGURE 10. Continuous line, variation in CO_2 concentration of the Ringer film covering the smooth muscle of frog's stomach during a contraction, such as results from the pH curve experimentally determined (log $CO_2 = 4\cdot56 - pH$). a, b, c, d, component phases, whose resultant is given by the continuous line. I, isometric contraction; St, stimulation.

ATP, although if it were ATP, it should be so, and at present we do not know any other substance the splitting of which could have these consequences. If it is really due to ATP, we must admit that the splitting of this substance is not a relaxation phenomenon, not even a shortening event but a pre-shortening phenomenon, at least in a smooth muscle.

In tetanized striated muscles we have found qualitatively the same four changes as in smooth muscle (figure 11). After a single shock the shortening, even at 0° C, is too fast to permit a true picture of the pH changes. We tried to use special needle glass electrodes inserted inside the muscle. But in this case it is impossible to get a steady pH at rest, due to injuries of the tissue; and the contraction, even in isometric conditions, produces slipping of the electrode and no reproducible pictures could be taken. The records on striated muscle have drawn our attention to the important fact that the initial length of the muscle has a very great influence on the shape of the *first alkaline* change which starts very early after stimulation. This was the starting point of our work on the pH changes due to the lengthening of the muscle at rest.

When a striated muscle at rest is stretched, it becomes alkaline (Dubuisson 1940, and figure 12). This was shown first by Margaria (1934). At the same time, the resting potential increases (Jacob 1942) and a quantitative relation exists between

the two phenomena which are perfectly reversible. We know also that there exist, when one stretches a muscle, an increase of double refraction (Fischer 1936), a decrease of volume (Fischer 1936, 1938) and an increase of impedance (Katz 1934). We had no idea at that time of the causes of these facts, but when one stretches a muscle one evidently distorts some elastic structures built from protein
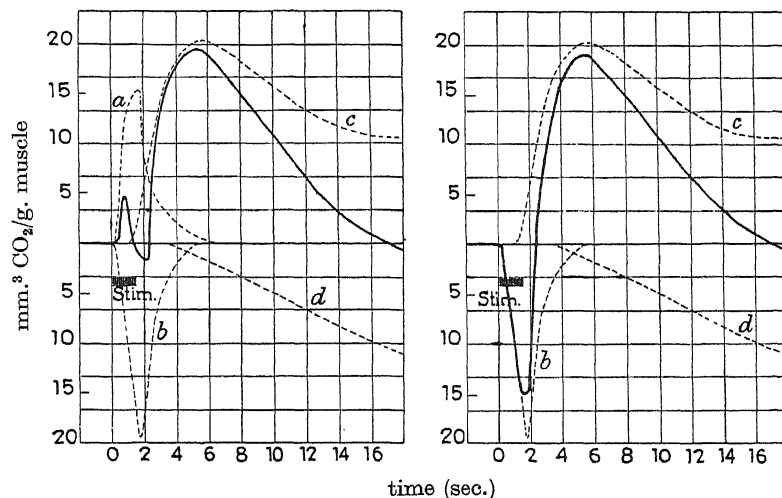
FIGURE 11. pH changes (full lines) of a frog's sartorius. *Left*, muscle fixed at 105 % of its initial length, isometric contraction; *right*, the same but isotonic contraction; dotted lines, component phases, whose resultant is given by the continuous line. *a* wave, isoelectric change (?) of muscle proteins; *b* wave, ATP splitting and resynthesis (?); *c* wave, phosphagen splitting and resynthesis; *d* wave, lactic-acid formation.
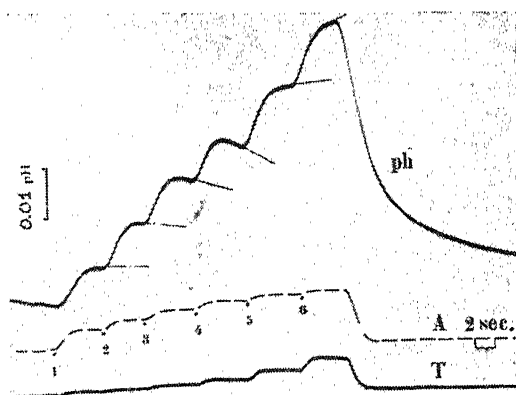
FIGURE 12. pH changes in frog's sartorius by stretching. Initial length of the muscle, 40 mm. In A, length of the muscle; T, muscle tension.

complexes, and there were some reasons to believe that the structure of those complexes could be more or less modified by the mechanical distortion *owing to the presence of weak bonds*, the breakage of which could give an uptake of H^+ ions and have many other chemical and physical consequences.

That was the point we reached in 1942 and so we started with the study of the muscle proteins and the complexes in which those proteins could be involved.

(c) *Protein changes*

I have no time to give you all the results of our work (Dubuisson, 1942 to 1950*b*; Jacob 1946 to 1948), but I should like to emphasize two points of it, because they may have a certain importance in a discussion upon chemical and physical affairs in muscle.

group I	group II	group III
l_1-l_2 myogens A and B	k_2 phosphorylase a phosphorylase b	h myoalbumin (Bate Smith)
$n-m$ aldolase hexosephosphate isomerase?	α, β, γ myosins globulin X (Weber)	
globulin X (Weber)		

$$10^{-5} \text{ cm.}/V/\text{sec.}$$
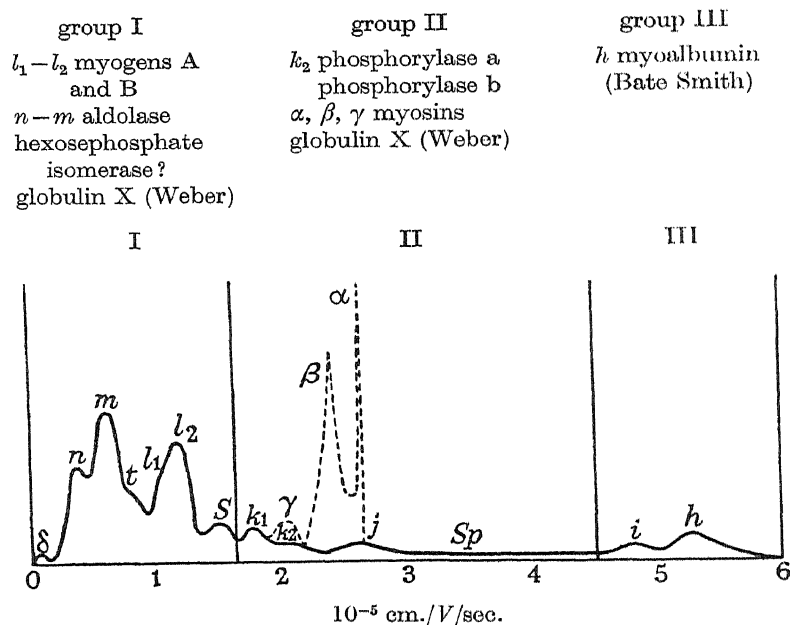
FIGURE 13. Electrophoresis diagram (Tiselius-Longsworth) of an extract of the rabbit muscle. Full line, extraction at μ 0·15; dotted line, new components appearing when the extraction is made at μ 0·40.

The first is this: if we extract the proteins of a muscle with salt solutions, we get in the Tiselius apparatus the picture in which the peaks α and β correspond to actomyosin and β myosins (figure 13). If, under the same conditions, we extract a muscle which has been exhausted by intensive stimulation, it is almost impossible to extract actomyosin, and the β myosins are very much less soluble (figure 14). The distribution of the other proteins is not modified.

More interesting are the facts when a fresh muscle is compared with a contracted one (for example by monoiodoacetic acid, by strychnine, by rigor mortis). In these cases, the β myosins are absolutely inextractable, a small quantity of actomyosin goes into the solution, and there appears a new component which we have called contractin.

There is no doubt that these phenomena result from the breaking or formation of links between the different elements of the complexes concerned in contraction and, in this sense, this kind of investigation is able to throw some light upon the mechanism of contraction and relaxation.

The second point is this: Straub (1942, 1943) has shown that it is possible to extract from muscle stroma, with rather drastic means, a protein which he has called *actin* and which may exist in two forms: *G-actin* and a polymer of this,

F-actin. If we study the electrophoretic behaviour of these two proteins, we see that the polymer (which has a very asymmetric body) moves faster in the electric field, which means that it carries more negative, or less positive, charge than the monomer. This is very surprising and can only be explained by supposing that the isoelectric point of the protein is very much lower. We could confirm these results by another method, using the Warburg apparatus: polymerization by K, Na, Mg
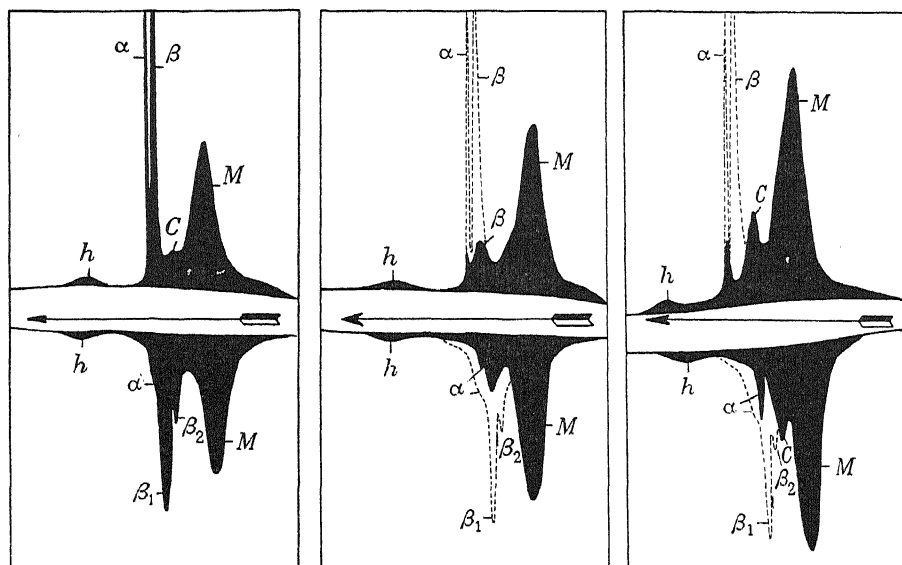
FIGURE 14. Electrophoresis diagram of muscle extracts of the rabbit, pH, 7·40, μ 0·35, $\sim$ 50·000 sec. electrophoresis. *Left*, normal muscle at *rest*; *middle, fatigued* muscle; *right, contracted* muscle (by stimulation and quick immersion in liquid air). *Above:* ascending boundaries; *below:* descending boundaries. In dotted lines, α and β peaks of the normal muscle at rest.

or Ca ions is accompanied by a decrease of pH. This is, I suppose, an important point because it shows that besides the well-known metabolic changes of ATP, phosphocreatine and glucose, some muscle proteins are able to give up or take off H^+ ions when there is a change in the state of polymerization. The pH increase found when a muscle is stretched could perhaps be explained, *in vivo* and *in situ*, by reversible depolymerization of actin or complexes of actin.

To sum up, the relations between physical and chemical events during muscle contraction and relaxation are made all the more difficult to correlate by the fact that we know only a very small part of the biochemistry of the proteins involved in the contraction cycle.

By the use of gentle methods, we can attempt to break up protein complexes stepwise by means of appropriate salt solutions, and by comparison of muscle in different functional states it is possible, by electrophoretic analysis, for example, to see that there are formations and breakages of bonds between proteins concerned in contraction and relaxation. This means that besides the metabolic changes of energetic substances, *the muscle machine itself undergoes chemical and physical*

changes, that there exists a real protein metabolism which must have its repercussion on many physical properties of the muscle, not only on its elastic and optical properties, but also on impedance changes, pH changes, heat production, double refraction, etc. Before we can correlate physical and chemical events with success, we should know not only the properties of the muscle proteins in solution, the investigation of which has given so many interesting facts by the hands of Weber, Edsall, Needham, Bailey, Szent-Györgyi and others, but chiefly the changes of the linkages of the proteins in the muscle machine itself during its activity, and there is very hard work to do in this field.

The nature of the fibril

By K. Bailey

One of the chief problems in muscle biochemistry at the present time is to elucidate the way in which the component proteins of the muscle fibril are assembled to give a contractile element. The myosin-actin-ATP system is so unusual that it is generally thought to contain the essential components for contraction, but it is true to say that *in vitro* studies of this system are merely suggestive, and have generally led to speculations on the mechanism of contraction which are either dubious or not amenable to experimental proof. The first difficulty is that in the molecular sense the nature of the interaction of myosin and actin is not clear. Under the conditions of salt concentration at which actin and myosin separately are dispersed in solution, the effect of ATP on actomyosin is to produce a large decrease in viscosity. This has been interpreted as a dissolution of the links which bind actin and myosin together, liberating the component molecules from an anastomosed network of fibres. More specifically, it seems due to the competition of ATP for certain groupings which not only link up with actin, but are responsible for the adenosine-triphosphatase activity of myosin. In salt-poor medium, where myosin by itself is no longer soluble, the effect of ATP is to produce an intense synaeresis, which superficially is suggestive of 'contraction'. There is no reason to suppose that the action of ATP in this case is at all different from that in a salt-rich medium; the breaking of the actin-myosin links is followed, however, by a further interaction which is reminiscent of the synaeresis of a fibrin clot. Jordan & Oster (1948) have concluded from light-scattering studies that, even at high ionic strength, ATP causes not a dissociation of actin and myosin, but an increased coiling of the actomyosin particle. There is no doubt, however, that the mechanical properties of actomyosin fibres in the presence of ATP are more in keeping with the original explanation that linkages responsible for anastomosis are broken down. What bearing the synaeresis phenomenon has upon the mechanism of contraction is not very clear, since the changes involved are inter- and not intramolecular. It is possible that the interaction reflects the means by which these water-holding proteins are tightly packed in the fibril.

These comments serve to show that we are not yet in a position to reconstruct the muscle fibril. Other difficulties arise in considering changes which occur in the

muscle fibre as a result of mincing. When a living muscle is minced, ATP breakdown is extremely rapid. At first, the sarcoplasmic fraction cannot be squeezed out by any moderate pressure, but soon, as glycolysis proceeds, the fibrils give up their water on simple squeezing by hand. The change is not wholly due to a fall in pH, which in certain physiological states can be made quite small. At this point also the extractability of fibrillar protein is much diminished. How are we to explain these facts? The synaeresis of fibre debris simulates that of actomyosin at low salt concentration in presence of ATP, although the overall ionic strength in muscle lies between that of the two systems I have described. Is the ATP inaccessible to the actomyosin of intact relaxed muscle, and are the changes which occur in the fibrillar protein on mincing induced by ATP during the few minutes it exists as such? It may be fruitful to attempt to relate the physical state of the fibrillar proteins of homogenized muscle with the appearance or disappearance of metabolites. This can be done by retarding breakdown with various enzyme poisons. So far, it has not been possible to prevent the hydrolytic breakdown of ATP, but it has been possible to poison both the triosephosphate dehydrogenase and the creatinephosphokinase.

This line of inquiry suggests that one might try to extend the Lundsgaard experiment on alactic contraction, since it might be possible to poison whole muscle so that the anaerobic resynthesis of ATP is entirely prevented. Direct changes in ATP content after repeated stimulation could be followed, and this procedure would have the advantage of measuring large metabolite changes, as distinct from the minute one which is likely to occur with a single twitch. It would not, however, distinguish events in contraction from those in relaxation. Unfortunately, the poisons which have produced this effect on mince have so far had no effect on whole muscle, due to the impermeability of the fibre membrane. Even if they entered the fibre, they might not produce the same effect on the enzyme systems of living muscle as on mince, and they might have other effects on the muscle proteins unfavourable to the maintenance of an active physiological state.

THE MOLECULAR MECHANISM OF CONTRACTION

BY M. G. M. PRYOR

From the physical point of view, muscle is an engine for the conversion of chemical energy into mechanical work, operating by means of reversible elastic strains in a solid working substance, the muscle proteins. In a system of this kind, only two types of cycle are possible, one involving changes in potential energy, and the other changes of entropy. In general, the evidence seems to be in favour of an entropy cycle, although there are some serious objections.

One important difference between the two types is that in an entropy cycle the strength of intermolecular attractions should decrease during contraction, and in a potential energy cycle they ought to increase; observations on whole muscles suggest that they do in fact decrease. The action of ATP on muscle proteins *in vitro*

agrees with this; as Professor Weber has remarked, it reduces the viscosity of solutions and liquefies gels. The most important objection to an entropy theory lies in the sign of the observed heat exchanges in active muscle. Professor Hill has pointed out that if the contraction of muscle were analogous to the contraction of stretched rubber, heat should be absorbed instead of being given out during shortening. I think this difficulty arises from the fact that to compare muscle to rubber implies that it functions as a heat engine, which is unlikely. It seems more probable that muscles convert chemical energy directly into work, and are capable of operating at constant temperature; if so, the type of entropy cycle we ought to consider is one which does not involve changes of temperature.

The theory of a 'chemical entropy engine' can be derived from a consideration of the properties of 'imperfect rubbers', in which the strength of the intermolecular attractions is not negligible. In an ideal rubber under strain the tension at constant length should increase with the absolute temperature in the same way as the pressure of a perfect gas at constant volume. This is true over a fair range of temperatures for some materials, but all natural and synthetic rubbers show a lower limit, the 'rubbery freezing-point', below which tension falls to a low value and is no longer proportional to the temperature; this behaviour is analogous to that of a vapour near its boiling point. At temperatures not far removed from the 'rubbery freezing-point', the tension in an imperfect rubber can be increased either by raising the temperature or by decreasing the strength of the intermolecular attractions by the addition of solvents or plasticizers. A model of such a system is afforded by the contraction of tendon in a solution of mercury-potassium iodide. The reagent (which is in fact neutral Nessler's solution) forms inert complexes with free amino groups, thus decreasing the number of polar linkages between molecules. Tendons immersed in the solution contract, and will exert considerable tension; the curve for isometric tension plotted against concentration is the same shape as the graph of tension plotted against temperature; at concentrations above the point of inflexion corresponding to the rubbery melting-point (A in figure 15), the tension bears an approximately linear relation to the concentration. This relation is reversible, so that an increase in tension tends to expel reagent from the tendon and a release to cause an uptake. If we suppose that in muscle a plasticizer is released on to the muscle proteins at the moment of stimulation, then there will first be an adsorption of plasticizer sufficient to melt the tendon, followed by a further uptake corresponding in amount to the extent to which the muscle shortens. If there is a positive heat of adsorption, Hill's heat of activation corresponds to heat of adsorption minus latent heat of 'melting', and his heat of shortening to heat of adsorption minus the heat absorbed in shortening. The heat exchanges would be difficult to measure by Hill's technique because of the large volumes of liquid which have to be used, but it is hoped to measure them calorimetrically.

The rate of shortening of tendons in mercury-potassium iodide bears the same relation to load as does the rate of shortening of muscle. As with muscles, the imposition of a load greater than the isometric tension causes a lengthening but, again as in muscles, the rate of lengthening is less than the value calculated from the equation for the rate of shortening. If both processes involve diffusion of

reagent, this difference is to be expected, as conditions for diffusion inwards are different from those for diffusion outwards.

One respect in which the model differs from muscle is that in the return half of the cycle, reagent is removed by dilution; to repeat the cycle with the same materials, work would have to be done in concentrating the solution. In muscle
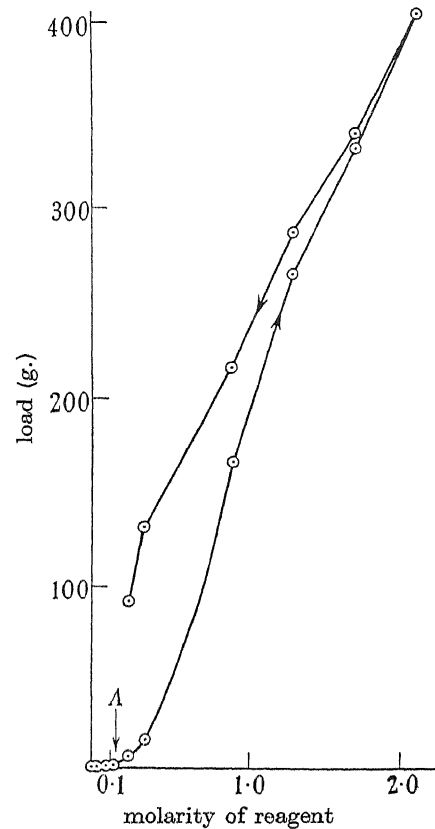
FIGURE 15. Variation of isometric tension with concentration of reagent for pheasant tendon in K_2HgI_2 solution. Load of 400 g. is approximately 80 kg./cm.^{2}.

the plasticizer (let us suppose ATP) is removed by chemical decomposition, and to repeat the cycle work has to be done to resynthesize it. On the analogy of the Carnot cycle we may expect the efficiency of the process to depend on the difference between the two working concentrations of plasticizer divided by the lower one, and if, as a result of the chemical reaction, the lower concentration is nearly zero, the efficiency might be correspondingly high.

THE ATP CONTENT OF RESTING AND ACTIVE MUSCLE

BY E. LUNDSGAARD

It is quite natural that the question whether ATP is broken down in a muscle twitch, especially in connexion with the contractile phase, has become a rather central point in this discussion.

In his introduction A. V. Hill mentioned that at present we do not even know if ATP under any conditions is broken down in the intact muscle. Undoubtedly A. V. Hill has deliberately exaggerated his statement on this point to some degree, probably to apply a kind of shock therapy to the biochemists interested in muscular contraction.

I do not think it can be doubted that ATP constantly is broken down and rebuilt in the intact muscle. Experiments with labelled phosphate demonstrate this convincingly. It is true, that it has never been demonstrated that the rate of turnover of phosphate in ATP is increased by stimulation. This, however, is simply due to the fact that even in the resting muscle the rate of turnover is so rapid that it keeps pace with the transfer of labelled phosphate from the surrounding fluids to the interior of the muscle cells. As ATP is the universal phosphate-donor I have no doubt myself that the increased rate of metabolism which accompanies a muscle twitch must involve an increase in the turnover of ATP-phosphate. But even if so, the breakdown of ATP might very well have to be ranged amongst the processes of recovery.

I know that it is going to be very difficult for the biochemists to satisfy A. V. Hill, for what he wants is a method which enables us to trace the time relations between mechanical response and breakdown of ATP. This of course is a formidable task, to use A. V. Hill's own expression. As long as no such method is available, however, we have to content ourselves with simpler methods.

I have carried out some determinations with a view to the question here discussed, using indeed very simple methods.

As will probably be remembered, Embden always claimed that a muscle frozen in liquid air is a contracted muscle. This claim dates back to 1922. I must confess that I never paid much attention to this claim of Embden, though I have seen dozens of times that a muscle frozen in liquid air attached to an isometric lever develops a considerable tension.

Shortly after I had received a copy of A.V. Hill's introductory remarks to this meeting to-day I carried out some experiments on isolated hind-limb preparations of cats poisoned with iodo-acetate. The purpose of these experiments was to investigate the phosphate exchange between plasma and muscle after poisoning with iodo-acetate. In these experiments one anterior tibial muscle was cut out and frozen in liquid air before the poisoning, the other anterior tibial muscle after the development of complete rigor. A most striking difference in the appearance of these two muscles reminded me of Embden's claim, that a muscle frozen in liquid air is a contracted muscle. The muscle cut out before the poisoning was quite short and thick and really looked 'contracted', whereas the muscle frozen after rigor had developed and consequently in an inexcitable state kept its long and slender resting form.

In this way reminded of Embden's maxim I thought it worth while to compare the content of ATP in muscles frozen in liquid air in an excitable state and muscle frozen down slowly in a deep-freeze at $-10°$ C and then frozen in liquid air.

As only slight differences in the content of ATP could be expected the determinations were carried out in the following way.

The muscles were ground while kept frozen* and extracted with trichloracetic acid. The extract was left overnight at 30° C to allow the creatine phosphoric acid to break down. The next day the inorganic P and the phosphocreatine P were precipitated and the samples left overnight at low temperature to secure complete precipitation. An aliquot fraction of the filtrate was then hydrolyzed with 1 N hydrochloric acid for 10 min. and the free phosphate determined. In this way the total amount of phosphate determined is ATP-phosphate, and even small differences should be detectable.

First, I carried out a series of determinations on muscles from temporaria. The frogs were cooled in ice-water. The two gastrocnemii were cut free, the tendon being cut as close to the muscle as possible. The muscles then were weighed separately and one thrown into liquid air, the other placed at − 10 °C for 2½ hr. After the lapse of this time, the second muscle, now stiff and inexcitable, also was thrown into liquid air. The results of this series are shown in table 2. The frog material at my disposal was rather poor which explains the rather great differences in ATP content from animal to animal. If the results are treated statistically as 'pairs', which of course is the only right thing to do, the probability that the difference between the two series is a random error is less than one per thousand.

Another series was carried out on gastrocnemii from esculenta which were in a somewhat better condition than the temporaria. In this series I tried to keep the muscle stimulated with a short tetanus while it was immersed in liquid air. I used needle-form electrodes thrusted into the belly of the muscle. The symmetrical muscle was stimulated as far as possible in the same way and frozen down slowly. This series was carried out at a temperature of about 5° C.

A. V. Hill mentioned in his introduction that nothing is gained by repeated stimulation when one wants to study the chemical changes occurring in connexion with the contractile phase. This of course holds true so far as such changes are concerned as are known to vary quantitatively with the degree of stimulation. If, however, one is looking for a qualitative difference between a contracted and a resting muscle it might be a help to try to keep the muscle in a stimulated state while it is fixed by freezing in liquid air.

The results of this series are shown in table 3. The average difference in ATP content between the two columns is somewhat greater than in table 2. The significance of the difference between the two columns is of the same order of magnitude as in table 1.

In this series it was necessary to freeze the muscles in liquid air while attached to the bone. For this reason the muscles had to be weighed after they were frozen. A systematic error might be introduced hereby, due to a loss of water from those muscles which are frozen down slowly before being thrown into liquid air. The average weight of the set of muscles frozen down slowly, however, was only 0·5 % lower than the average weight of the set frozen directly, whereas the average difference in ATP is 8·5 %.

---

* To avoid loss of material the muscles were wrapped in a small piece of gauze while being ground.

A. V. Hill and others

Whether the greater part of a muscle frozen in liquid air in an excitable state is fixed during contraction or during relaxation we do not know. Crude experiments of this kind I have presented here can of course not satisfy A. V. Hill. They seem, however, to indicate that ATP actually is broken down during the contraction-relaxation cycle. Whether ATP is broken down during contraction and rebuilt

TABLE 2. ATP CONTENT IN MG. P%

I	II	diff.	I	II	diff.
14.7	17·3	+2·6	15·2	14·9	−0·3
13·3	14·7	+1·4	16·1	18·5	+2·4
13·7	13·6	−0·1	14·5	14·8	+0·3
17·2	18·2	+1·0	14·8	15·8	+1·0
16·6	14·9	−1·7	14·7	15·6	+0·9
18·2	19·8	+1·6	16·4	18·4	+1·8
15·4	17·5	+2·1	14·3	14·3	0·0
14·3	16·2	+1·9	15·6	17·0	+1·4
13·3	13·7	+0·4	14·9	14·3	−0·6
14·5	15·3	+0·8	13·5	12·9	−0·6
20·5	21·8	+1·3	13·6	14·8	+1·2
15·3	16·2	+0·9	16·5	17·4	+0·9

	I	II	diff.
average	15·296	16·154	0·858

TABLE 3.

I	II	diff.	I	II	diff.
17·7	19·5	+1·8	22·2	23·8	+1·6
24·4	25·1	+0·7	19·2	19·7	+0·5
24·4	24·9	+0·5	19·4	24·4	+5·0
23·6	25·2	+1·6	15·1	14·7	−0·4
25·2	28·4	+3·2	15·2	18·1	+2·9
24·4	28·2	+3·8	19·8	20·3	+0·5
16·9	21·4	+4·5	18·5	20·3	+1·8
20·7	20·4	−0·3	19·8	20·4	+0·6
18·9	21·1	+1·2	19·0	19·8	+0·8
14·8	17·2	+2·4			

	I	II	diff.
average	19·96	21·68	1·72

during relaxation, or whether ATP is broken down and rebuilt during relaxation one cannot assume that relaxation is chemically 'neutral' as it is thermally 'neutral'.

Though only limited information can be gained through such simple experiments as mine, I find my preliminary results sufficiently encouraging to make it worth while to repeat the experiments. They may at any rate be of some technical significance by reviving Embden's maxim, that freezing a muscle in liquid air produces chemical changes in it, for which reason this procedure cannot be used when one wants to fix a muscle in a definite mechanical and chemical state.

ATP BREAKDOWN AND MUSCLE CONTRACTION

By DOROTHY NEEDHAM, F.R.S.

I should like to consider first the possibility of showing, by chemical methods, that dephosphorylation of ATP is associated with normal contraction in a single twitch or a short series of twitches. There is some hope that this could be done, using the enzymatic spectrophotometric methods, worked out by Kalckar (1947) for purine analysis: the experiments would, of course, have to be done at 0° C, and on slow-moving muscle as Professor Hill has suggested.

In visualizing the following method I am greatly indebted to Professor Kalckar for his advice. The intense absorption in the ultra-violet, which is characteristic of the purines, occurs always in the region between 250 and 290 mμ, but the absorption maxima vary in height and in position for the different compounds. The method suggested for attacking this problem is to follow the production of ADP. ADP has the same absorption spectrum as ATP; but if it is acted upon by purified myokinase and purified adenylic deaminase, it is converted to the extent of 50 % into inosinic acid:

$$2ADP \underset{}{\overset{\text{myokinase}}{\rightleftharpoons}} ATP + AA$$

$$AA \xrightarrow{\text{AA deaminase}} \text{inosinic acid} + NH_3$$
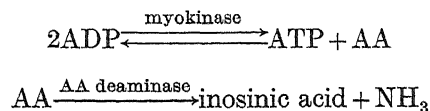
The inosinic acid is already spectrophotometrically distinguishable from ATP, having its absorption maximum at 240 mμ instead of 265; but with the small quantities of ADP to be expected in these experiments better differentiation is needed. This is to be obtained by hydrolyzing the inosinic acid to set free hypoxanthine and after correct adjustment of the conditions, oxidizing the hypoxanthine with xanthine oxidase. The uric acid quantitatively resulting can be determined by its absorption at 293 mμ—a maximum now well removed from that of the adenylic compounds. This method has the great advantage that it aims at estimating a small increase from zero concentration (since there is probably no ADP in the resting muscle).

According to the heat measurements, if the heat production is dependent on ATP dephosphorylation, one expects about 8 μg. inorganic P to be set free in a twitch by 1 g. muscle, or about 1·5 μg. for a 200 mg. muscle. This would correspond to about 3·2 μg. hypoxanthine (since only 50 % of the ADP is converted into hypoxanthine)—about 1 μg./ml. final concentration under the conditions of the experiment. It should be possible to estimate this with about 5 % accuracy.

The possibility of determining the full ATP breakdown associated with single twitch would depend on at least two conditions: (a) In order to extract the purine compounds from the stimulated muscle and the resting control and obtain them in solution fit for spectrophotometric analysis, the muscles must be extracted with perchloric acid of a strength to denature the proteins. The muscles could be crushed rapidly at 0° C in the extractant but we do not yet know how rapidly penetration and consequent cessation of enzyme activity would take place, nor do we know what effect the traumatic stimulus would have on ATP breakdown, and how it would

compare quantitatively with the effect of the twitch itself. Instantaneous freezing of the muscles in liquid air before extraction is open to the criticism that the cold itself acts as a strong stimulus and even a resting muscle is frozen in the contracted state. (b) According to current theory the ADP formed is rephosphorylated by interaction with creatine phosphate and we do not know how rapidly this takes place after a single twitch. It might be so rapid as to obscure part of the breakdown.

Using such a method, one might hope to find at least whether ATP breakdown does go on in normal, unfatigued contraction. With luck, one might get some idea of the amount breaking down in a single twitch; should, however, the breakdown due to trauma (occurring also in the unstimulated control) prove too high, one might hope to build up a definite difference between the stimulated and the resting muscle by several consecutive twitches. It is interesting in this connexion that Lundsgaard found, with IAA-poisoned muscles at 0° C, that creatinephosphate breakdown continues after relaxation from a 25 sec. tetanus, half as much again being formed after the end of relaxation. This gives us some grounds to hope that, in the normal, unpoisoned muscle too, creatinephosphate rephosphorylation of ADP does not take place so rapidly as to make ATP breakdown indistinguishable.

The possibility also exists that, by comparing the ADP content immediately after relaxation with that after a suitable interval of time, light might be thrown on the question as to whether ATP breakdown is confined to the contraction-relaxation period or is to be regarded as a recovery process; there seems at present, on the other hand, little hope of distinguishing by such biochemical methods, whether the ATP breakdown is associated with the contraction phase or the relaxation phase.

In connexion with the general discussion as to whether ATP breakdown (if it can be proven) is one more recovery process or is really associated with the contraction-relaxation period, I think it may be said that the more convincing the evidence that myosin or actomyosin itself is ATPase, the more likely is it that the ATP breakdown is actually associated with the changes in form of the muscle fibrils and not merely with the energetics of recovery. Bailey & Perry (1947) have recently obtained very convincing evidence that ATP does actually combine with certain SH groups of the myosin molecule.

The fact that classically prepared myosin preparations, used in vitro as ATPase, have a pH optimum above 9 and show very little activity around pH 7·4, together with the fact that the Ca activation of the enzyme is antagonized by Mg when the two ions are present together in proportions such as exist in the muscle, has been taken sometimes to mean that myosin cannot in vivo be acting as ATPase. However, the recent work of Keilley & Meyerhof (1948) shows that actomyosin (as distinct from myosin) can have a pH optimum for ATPase activity at about 7·5, activated either by Ca or Mg. More work is needed here.

Finally, I should like to make one comment about the relaxation phase, where no heat is given out (unless work is degraded into heat). There is discussion as to whether, in view of this, the relaxation phase could be the period of any important chemical reactions. I personally find it quite conceivable that during this period, such a reaction as ATP breakdown may be going on and bringing about, for

example, reconstitution of certain protein linkages with 100 % thermal efficiency. In emphasizing this possibility, I would point to the anaerobic recovery period, occupying 3 or 4 min. after contraction and relaxation under anaerobic conditions, during which it is known that glycogen breaks down to lactic acid, little or no heat is given out and creatinephosphate is formed. The free energy required for this synthesis is provided by the glycogen breakdown; some dozen chemical reactions are involved yet the thermal efficiency is over 90 %.

MODEL OF A FOLDED POLYPEPTIDE CHAIN

BY A. ELLIOTT

The models on view in the ante-room show a way of folding a polypeptide chain which is consistent with some observations we have recently made with polarized infra-red radiation (Ambrose & Hanby 1949; Ambrose, Elliott & Temple 1949). The α-folded proteins, keratin, myosin and tropomyosin, have been found when oriented to show greater absorption of the N-H frequency when the electric vector of the absorbed radiation is in the direction of the fibre axis, hence the N-H bond must be preferentially oriented in this direction. A study of models has suggested that the only likely folding of the polypeptide chain consistent with this fact involves a seven-membered ring containing two amino-acid residues; the ring is completed by hydrogen bonds:
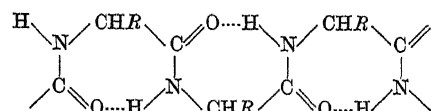
A new type of atomic model which has been developed in our laboratories has been used. The scale is 0·8 in. to the Angstrom unit. The valency links, while allowing free rotation about single co-valent bonds, also allow some distortion of the bond angles when strains occur but are strong enough to allow long polymer chains to be built. The molecular model exhibited shows twenty-four amino-acid residues, with side chains on one side of the back-bone, representative of those occurring in myosin; the side chains on the other side have been removed for clearness and their positions indicated by single carbon atoms.

The model, arranged in the proposed α-fold, is suspended from a curtain rail bent parallel to itself. This allows it to be seen that the secondary folding of the proposed structure can occur through free rotation about single co-valent bonds without necessitating distortion of bond angles and can be close enough for side chains on either side of the bond to come into contact. This would allow salt-linking between acid and basic groups. The model is incomplete and no attempt is made to explain fully the mechanism of folding. It is, however, perhaps of interest to point out that if such secondary folding of myosin molecules does occur during the contraction of muscle, the number of interacting groups would be proportional to the shortening, and that this might be correlated with A. V. Hill's observations that the heat of shortening is proportional to the shortening.

Ultra-violet and infra-red investigations on muscle

By R. Barer

Infra-red absorption spectroscopy of muscle has already been carried out, using the Burch reflecting microscope (Barer, Cole & Thompson 1949: Barer, Thompson & Williams unpublished). There are considerable difficulties involved in this type of work. In the first place it is rather doubtful whether such measurements will ever be possible on living muscle owing to the presence of water, which possesses intense absorption bands in some of the most useful regions of the infra-red spectrum. It may be possible to overcome this difficulty to some extent by using heavy water which has a different absorption spectrum. It is in principle possible to obtain information similar to that given by infra-red spectroscopy, even in the presence of water, by means of Raman spectroscopy, but the technical difficulties involved, particularly light scattering by colloids, would seem to preclude this method of attack so far as muscle is concerned. Our infra-red measurements have hitherto been confined to dried material. The results indicate that there is little prospect of working with whole muscles, as even single isolated striated fibres of the frog, rabbit and crab were usually too thick. However, it was possible to obtain good spectra in the chemically important region from 3 to 14 μ, on exceptionally thin single fibres or on artificially compressed fibres. An attempt was made to detect dichroism by means of polarized infra-red radiation, but to our surprise none was observed throughout the 3 to 14 μ range, even though the material used showed strong birefringence in the visible region. Dr Stocken and I have recently examined certain molecular models of muscle, in the light of the work of Ambrose, Elliott & Temple (1949) on myosin, and it now appears possible that infra-red dichroism of muscle might be expected to manifest itself only under rather special conditions. We hope to put these theoretical deductions to experimental test.

As regards measurements on muscle in the ultra-violet region, the position is much more promising. It is quite possible to determine the absorption spectrum of the A or I band in living single fibres. The entire spectrum from about 230 mμ in the ultra-violet to over 600 mμ in the visible can be recorded simultaneously, using the reflecting microscope. This technique can also be used with polarized ultra-violet light, in order to detect variation of dichroism in crystals at different wave-lengths (Barer, Jope & Perutz unpublished), and I intend to apply it to the study of dichroism in muscle fibres. Another new possibility is the observation of birefringence, as well as dichroism, in the ultra-violet. I have recently carried out experiments with a view to developing a new type of ultra-violet polarizer and it should now be possible to use the reflecting microscope as an ultra-violet polarizing microscope.

Preliminary work will be carried out on sections and on resting single fibres, but if the methods prove successful an attempt will be made to investigate the events occurring during contraction by using a photo-multiplier tube in conjunction with a cathode-ray oscillograph. The speed and sensitivity of such electronic methods should make it possible to record variations in absorption spectrum, dichroism and birefringence in the A and I bands, even during a single twitch.

By M. G. M. Pryor

Professor Astbury has said that a contracted muscle is like steel, but this is perhaps not literally true; in fact its mechanical properties are more like rubber. Resting muscle or hair behaves mechanically like 'imperfect rubbers' below their rubbery freezing-point. When a muscle contracts or a hair 'supercontracts' the relaxation time is decreased, the elastic modulus (as measured at a moderate rate of strain) is decreased, and the material behaves more like an ideal rubber: these changes are most easily explained by a decrease in the strength of intermolecular attractions.

By I. MacArthur

I gather Dr Pryor holds the term 'steel-like' or 'solid' should be less appropriate to contracted than to relaxed muscle, and inclines to the entropy school. While the respective entropy and energy contributions in elasticity vary with nature and conditions (mobility, cross-linking, swelling, ease of molecular packing, nature of potential fields, concentration of polar groups, etc.), relevant measurements on fibrous proteins in fact confirm rather the importance of energy; further, super-contraction in keratin, usually an irreversible change following cross-link rupture, can occur reversibly using cuprammonium.

The answer to Professor Hill's query as to whether the fine structure of muscle during stimulation is determinable by X-ray diffraction methods with existing or envisaged equipment, involves such factors as utilizable X-ray power, recording technique, specimen stability and nature of reflected spectrum. Our highest-power X-ray generator gives too strong reflexions from a *single crystal* (e.g. dibenzil) even limited to a single rotation on a Unicam spectrometer. Using merely additional engineering and electronic expedients, many reflexions could be recorded, precisely in position and to 5 % accuracy in intensity, in some 7 msec. This seems promising until we recall experience with keratin and dried muscle and the nature of the probable spectrum: a fibrous pattern, low reflectivity, diffuse lines in at least one direction, macro-patterns involving close spots and therefore high-resolution fine beams and preferably monochromatic rays, key central lines possibly swamped by diffuse small-angle scattering, and additional complications. Useful fibre X-radio-grams (ramie) have been recorded in 7·5 sec., and the contractile cycle during ion-exchange in swollen fibrous alginates in 1/2 to 3/4 min., but these are simpler problems. Powerful 'flash' tubes (2000 amp. for 6 μsec.) are excellent for phase selection, but yield the equivalent output of $\sim$1 sec. continuous exposure of a normal commercial sealed-off tube, with very moderate focus and limited life. The major problem is muscle deterioration; even the heating effects of large bursts of X-rays are by no means negligible.

Thus the prospect reduces to repetitive recording of the same phase (by syn-chronizing stimulus and exposure) at the best on one muscle, if necessary on relays of muscles sufficiently similar. The most probable solution will come through greatly increased sensitivity of recording, either by successive spot analysis by

Geiger counter, or areally on the lines of television scanning. Another possibility is high-voltage electron diffraction, the specimen being vacuum-shielded by Be film.

Using the repetitive synchronized procedure, present technique is on the verge of feasibility; the importance of the result might well justify the time and cost of special preparations.

By J. T. Edsall

(At the meeting Professor Edsall gave a general summary of the remarks of the other speakers. For the printed account of the proceedings he has provided the following original contribution.)

Many authors have suggested analogies between the elastic properties of muscle and those of rubber; I should like to emphasize the structural characteristics of myosin which are most unlike those of the non-polar chains of rubber—the numerous ionic and polar groups. It is certainly significant that (see Bailey 1948) the proportion of charged and polar side-chain groups is higher in myosin than in any other protein yet analyzed, except tropomyosin and the protamines. We may express the concentrations of these groups as chemical equivalents per litre of solid myosin (assuming a density of approximately 1·3) and then divide by a factor of the order 5* to obtain the actual equivalent concentrations in the fibrils. The results are shown in table 4.

TABLE 4. EQUIVALENT CONCENTRATIONS OF ANIONIC, CATIONIC AND HYDROXYL GROUPS IN SOLID MYOSIN (DENSITY 1·3)

	equiv./l.
free anionic (carboxyl) groups	2·0
cationic groups:	
histidine	0·20
arginine	0·55
lysine	1·05
hydroxyl groups:	
serine	0·54
threonine	0·56
tyrosine	0·25

Values for anionic groups from titration data of Dubuisson & Hamoir (1943); histidine, arginine and lysine from Macpherson (1946); hydroxyamino-acids from Bailey (1948).

The figures in table 4 are certainly large, even when divided by a factor of 5 or more, which makes them of the same order of magnitude as that of the free ions within the muscle fibre (see Dubuisson 1942). The carboxyl groups should all be negatively charged at any pH near 7; the arginine and lysine residues should all be

* Myosin makes up approximately 10 % of the total muscle weight. Its concentration within the fibrils may be of the order of 20 % or higher. Actin has been omitted from this calculation, the purpose of which is primarily illustrative, since it is quantitatively less important than myosin, and amino-acid composition data for actin appear to be lacking.

positively charged; the histidine groups should be positively charged or uncharged, in proportions varying markedly with pH in the neighbourhood of 7. As with most proteins, the total number of charged groups at pH 7 is vastly greater than the net charge. For many phenomena, the total charge is by far the more important of the two, as witnessed for instance by the selective binding of many anions by serum albumin, even when its net charge is negative (see Scatchard 1949 and Klotz & Urquhart 1949). The presence of positively charged groups on the protein appears to be a necessary, though certainly not a sufficient, condition for such binding.

Concerning the distribution of the charged groups of myosin in space we know nothing as yet; but the interionic forces they exert on one another, and on the surrounding ions, must be large. Anyone who has studied the solubility of proteins is aware of the profound changes in the free energy* of a protein molecule which are produced by moderate changes in the ionic strength or dielectric constant of the surrounding medium. In a compact rigid molecule, such as a small corpuscular protein, the shape remains essentially unchanged while the free energy changes; in a flexible structure changes of shape are to be expected as well.

Certainly the electrostatic forces between the charged groups in myosin must vary as the protein chains fold or unfold, whatever the driving mechanism of the contractile process. The presence of a net charge—positive or negative—on the chains tends to extend them by repulsion; the neutralization of the charge favours coiling. These expectations are strikingly verified in solutions of synthetic polyelectrolytes (see, for instance, Fuoss & Strauss 1949). Twenty years ago, indeed, such an electrostatic mechanism for muscular contraction was suggested by K. H. Meyer (1929). His attempt to explain the change in net charge as due to a variation in pH during contraction and relaxation was challenged by Weber (1930); the strength of Weber's arguments appeared decisive to most biochemists, and the idea received little consideration for many years. Recently it has been revived in a new form by Riseman & Kirkwood (1948), who suggest a charging of myosin by phosphorylation of the hydroxyl groups of serine and threonine residues by ATP. Their statement of the electrostatic hypothesis of contraction is probably the clearest yet given; discussion of their particular assumption concerning ATP is probably best deferred for the moment. At this conference, Hill has challenged the evidence that ATP has any particularly close relation to the contractile process; and several biochemists appear to be taking up that challenge. I should like to wait for their evidence before discussing further the function of ATP.

However, one may insist that the electrostatic factors are important in relation to the contractile process, even if they are not decisive; and no theory of the process can afford to ignore them. We may briefly summarize the major factors which may modify the electrostatic forces between the charged groups in the myosin fibres:

(1) Binding or release of protons due to pH changes. This was the principal factor invoked by Meyer. From what we know of protein titration curves, the imidazole groups of histidine residues, and any free α-amino groups that may exist at the ends

---

* More precisely the chemical potential, or partial molal free energy.

of peptide chains, are the most likely to be involved in such reactions. Such pH changes in muscle are probably not negligible, but are also not likely to be of decisive importance.

(2) Binding or release of other ions. With regard to ATP, or the phosphate ions derived from it, this idea has repeatedly been discussed. Snellmann & Erdös (1948) have found by electrophoresis evidence of strong binding of Ca^{++} to myosin, following earlier evidence produced by Szent-Györgyi (1947). However, it is very dubious whether calcium plays a significant physiological role. Binding of potassium by myosin has been much discussed, but reliable figures appear to be nearly non-existent. Dr Conway has referred, in the discussion at this conference, to evidence that the potassium ions in muscle are all—or nearly all—free in solution. There is urgent need for quantitative studies in this field, similar to those carried out by Scatchard, Scheinberg & Armstrong (1950) and his associates on the binding of anions (chloride and thiocyanate) by serum albumin.

(3) Variations of ionic strength of the medium surrounding the myosin fibre. These may, of course, be extremely localized, and will in any case be very difficult to detect, if transitory. If they occur, however, they are likely to be of profound importance; certainly such changes are decisive in determining the contracted or extended form of synthetic polyelectrolytes (see Fuoss & Strauss 1949).

(4) Variations of dielectric constant of the medium. These again may be of major importance, and there is some hope of experimental approach to the problem to-day. However, the macroscopic dielectric constant, measured on a system as heterogeneous as muscle, may not tell us much of the forces between charged groups in a small region within the fibre. In any case, the 'effective dielectric constant' (Kirkwood & Westheimer 1938) within the fibre is likely to be considerably lower than that of water, since the protein molecules themselves act like large cavities of low dielectric constant embedded in the aqueous medium. Moreover, myosin (actomyosin), as ordinarily prepared, contains considerable amounts of lipids (Greenstein & Edsall 1940). These are commonly removed by extraction before the protein is dried for analysis, but in the muscle fibre they probably exist in close association with myosin, and tend to lower the dielectric constant still further, besides exerting more specific chemical effects.

The electrostatic theory as here outlined faces at least one major difficulty. The extension of a protein chain by electrostatic repulsion due to formation of charged groups along the chain obviously requires the expenditure of free energy; it would certainly appear to be an active rather than a passive process. Can such a picture be reconciled with the finding of Hill (1949) that if chemical changes occur during relaxation their net thermal effect is negligibly small? This certainly points to the conclusion that relaxation is a passive process. It is, of course, possible that the net thermal effects are very small in relaxation, even though the free-energy changes are large, but one cannot feel happy about such a conclusion unless it is possible to propose a mechanism that explains it. It may be that Hill's findings will rule out the type of electrostatic mechanism I have discussed as a major factor in the contractile process; but for the moment I think we can only suspend judgement about this.

Certainly some of the other phenomena described by Hill at this conference furnish formidable problems for any chemist to interpret. The fact that all the work done in a slow stretch under a constant load is absorbed by the muscle practically drives us to the assumption that the process of lengthening leads to the reversal of chemical reactions which accompany shortening. This is to me certainly a most startling result; I think it should be possible even to-day to propose chemical models which would lead to such an effect, but they are likely to be far removed from the reality. The fact that heat production definitely precedes mechanical response, which appears to be conclusively demonstrated by the latest results in Hill's laboratory, is less startling; in fact, it seems inherently a most reasonable thing to expect, though not so easy to explain. Also the fact long suspected, but only recently proved, that the muscle fibre is set almost instantly by the stimulus into a maximum state of activity, is a fact of major importance, which any theory of the contractile process must reckon with.

I cannot refrain from a few remarks about the theory of rubber-like elasticity and its relation to muscle. The evidence leads me to suspect that the coiling of the myosin chains in muscle must be a very different kind of thing from the random type of coiling which occurs when a stretched piece of rubber is allowed to contract. The transition from relaxed to contracted muscle is presumably a transition from one type of order to another type of order, rather than from order to disorder. The degree of order in the contracted fibre may be less perfect than that in the relaxed, and large entropy changes may be involved in the contractile process. However, consideration of the X-ray data of Astbury and McArthur, imperfect as the information they give may be, does seem to me to be in better accord with the kind of view I have suggested. In saying this, of course, I have no intention of criticizing the kind of analysis undertaken by Dr Pryor, whose model studies on tendon may prove very illuminating indeed.

REFERENCES

Ambrose, E. J. & Hanby, W. E. 1949 *Nature*, **163**, 483.
Ambrose, E. J., Elliott, A. & Temple, R. B. 1949 *Nature*, **163**, 859.
Astbury, W. T. 1947*a* *Proc. Roy. Soc.* B, **134**, 303.
Astbury, W. T. 1947*b* *Nature*, **160**, 388.
Astbury, W. T. 1949 *Experimental cell research*, suppl. 1, p. 234 (being the published proceedings of the 6*th Int. Congr. Experimental Cytology, Stockholm*, 1947).
Astbury, W. T., Perry, S. V., Reed, R. & Spark, L. C. 1947 *Biochim. Biophys. Acta*, **1**, 379.
Astbury, W. T. & Weibull, C. 1949 *Nature*, **163**, 280.
Bailey, K. 1948 *Biochem. J.* **43**, 271.
Bailey, K. & Perry, S. V. 1947 *Biochim. Biophys. Acta*, **1**, 506.
Barer, R., Cole, A. R. H. & Thompson, H. W. 1949 *Nature*, **163**, 198.
Bear, R. S. 1945 *J. Amer. Chem. Soc.* **67**, 1625.
Buchthal, F., Deutsch, A. & Petersen, A. 1947 *Acta Physiol.* **13**, 167.
Buchthal, F. & Knappeis, G. G. 1940 *Scand. Arch. Anat. Physiol.* **83**, 281.
Cole, K. S. 1938 *XVI Int. Physiol. Congr. Zürich*.
Draper, M. H. & Hodge, A. J. 1949 *Nature*, **163**, 576.
Dubuisson, M. 1935 *Arch. int. Physiol.* **41**, 177.
Dubuisson, M. 1936 *C.R. Soc. Biol., Paris*, **122**, 817.
Dubuisson, M. 1937*a* *J. Physiol.* **89**, 132.
Dubuisson, M. 1937*b* *Proc. Soc. Exp. Biol., N.Y.*, **35**, 609.

Dubuisson, M. 1937c *J. Physiol.* 90, 47P.
Dubuisson, M. 1937d *Pflüg. Arch. ges. Physiol.* 239, 314.
Dubuisson, M. 1938a *XVI Int. Physiol. Congr. Zürich.*
Dubuisson, M. 1938b *Ann. Physiol.* 14, 522.
Dubuisson, M. 1939a *Liège Méd.* no. 6, 1.
Dubuisson, M. 1939b *J. Physiol.* 94, 461.
Dubuisson, M. 1939c *Ann. Physiol.* 15, 445.
Dubuisson, M. 1940 *Arch. int. Physiol.* 50, 203.
Dubuisson, M. 1942 *Arch. int. Physiol.* 52, 439.
Dubuisson, M. 1946a *Experientia,* 2, 258.
Dubuisson, M. 1946b *Experientia,* 2, 412.
Dubuisson, M. 1947a *C.R. Soc. Biol., Paris,* 141, 800.
Dubuisson, M. 1947b 'Différenciation électrophorétique et séparation de diverses composantes dans les myosines de muscles au repos et fatigués, de Mammifères et de Mollusques.' Un symposium sur les Protéines, Paris.
Dubuisson, M. 1947c *Algérie Méd.* p. 195.
Dubuisson, M. 1947d *Experientia,* 3, 213.
Dubuisson, M. 1947e *Exposés Annuels de Biochimie Médicale,* 9e série.
Dubuisson, M. 1947f *Congr. Int. Physiol. Oxford.*
Dubuisson, M. 1947g *Experientia,* 3, 372.
Dubuisson, M. 1947h *Bull. Acad. Sci. Belg.* 33, 769.
Dubuisson, M. 1948a *Experientia,* 4, 152.
Dubuisson, M. 1948b *Proc. 6th Int. Congr. Exptl. Cytology, Stockholm,* p. 257.
Dubuisson, M. 1948c *Arch. int. Physiol.* 61, 93.
Dubuisson, M. 1948d *Experientia,* 4, 437; *Bull. Soc. chim. Biol.* 31, 114.
Dubuisson, M. 1948e *Bull. Acad. Roy. Sci. Belg.* 34, 978.
Dubuisson, M. 1949 *Congrès Assoc. des Physiol. de langue française, Montpellier.*
Dubuisson, M. 1950a *Biol. Rev.* 25, 46.
Dubuisson, M. 1950b *Biochim. Biophys. Acta* 4, 25.
Dubuisson, M. & Debot, A. 1940 *Arch. int. Physiol.* 50, 54.
Dubuisson, M. & Hamoir, G. 1943 *Arch. int. Physiol.* 53, 308.
Dubuisson, M. & Jacob, J. 1945a *Bull. Soc. Roy. Sci. Liège,* no. 3, p. 133.
Dubuisson, M. & Jacob, J. 1945b *Bull. Soc. Roy. Sci. Liège,* no. 3, p. 145.
Dubuisson, M. & Jacob, J. 1945c *Rev. Can. Biol.* 4, 426.
Dubuisson, M. & Jacob, J. 1945d *Experientia,* 1, 272.
Dubuisson, M., Lecomte, J. & Monnier, A. M. 1942 *Arch. int. Physiol.* 52, 408.
Dubuisson, M. & Roubert, L. 1947 *C.R. Soc. Biol., Paris,* 141, 802.
Dubuisson, M. & Schulz, W. 1938 *Pflüg. Arch. ges. Physiol.* 239, 776.
Engelhardt, W. A. & Ljubimova, M. N. 1939 *Nature,* 145, 668; *Biochemija* (Russ.), 4, 716.
Fischer, E. 1936 *J. Cell. Comp. Physiol.* 8, 503.
Fischer, E. 1938 *XVI Int. Physiol. Congr. Zürich.*
Fuoss, R. M. & Strauss, U. P. 1949 *Ann. N.Y. Acad. Sci.* 51, 836.
Greenstein, J. P. & Edsall, J. T. 1940 *J. Biol. Chem.* 133, 397.
Hall, C. E., Jakus, M. A. & Schmitt, F. O. 1946 *Biol. Bull. Woods Hole,* 90, 32.
Hartree, W. 1933 *J. Physiol.* 79, 487.
Hill, A. V. 1938 *Proc. Roy. Soc.* B, 126, 136.
Hill, A. V. 1939 *Proc. Phys. Soc.* 51, 1.
Hill, A. V. 1949 *Proc. Roy. Soc.* B, 136, 211.
Hill, D. K. 1949 *J. Physiol.* 108, 292.
Jacob, J. 1942 *Arch. int. Physiol.* 52, 417.
Jacob, J. 1946a *Experientia,* 2, 110.
Jacob, J. 1946b *Biochem. J.* 41, 83.
Jacob, J. 1946c 'Classification et caractérisation des protéines musculaires.' Un symposium sur les Protéines, Paris.
Jacob, J. 1947 *Experientia,* 3, 241; *Congr. Int. Physiol. Oxford.*
Jacob, J. 1948 *Biochem. J.* 42, 71.
Jakus, M. A. & Hall, C. E. 1947 *J. Biol. Chem.* 167, 705.
Josenhans, W. 1949 *Z. Biol.* 103, 61.
Jordan, K. J. & Oster, G. 1948 *Science,* 108, 188.

Kalckar, H. M. 1947 *J. Biol. Chem.* **167**, 445.

Kamp, F. 1941 *Biochem. Z.* **307**, 226.

Katz, B. 1934 *Pflüg. Arch. ges. Physiol.* **234**, 510.

Keilley, W. W. & Meyerhof, O. 1948 *J. Biol. Chem.* **176**, 591.

Klotz, I. M. & Urquhart, J. M. 1949 *J. Amer. Chem. Soc.* **71**, 1597.

Kirkwood, J. G. & Westheimer, F. H. 1938 *J. Chem. Phys.* **6**, 506, 513.

MacArthur, I. 1943 *Nature,* **152**, 38.

MacClendon, G. F. 1929 *Amer. J. Physiol.* **91**, 83.

Macpherson, H. T. 1946 *Biochem. J.* **40**, 470.

Margaria, R. 1934 *J. Physiol.* **82**, 496.

Matoltsy, A. G. & Gerendas, M. 1948 *Hung. Acta Physiol.* **1**, 116.

Meyer, K. H. 1929 *Biochem. Z.* **214**, 1.

Muralt, A. von 1935 *Ergebn. Physiol.* **37**, 406.

Needham, D. M., Needham, I., Shen, S. C. & Lawrence, A. S. C. 1941 *Nature,* **147**, 776.

Polis, D. B. & Meyerhof, O. 1947 *J. Biol. Chem.* **169**, 389.

Portzehl, H. 1950 *Z. Naturf.* (in the Press).

Portzehl, H., Schramm, G. & Weber, H. H. 1950 *Z. Naturf.* (in the Press).

Riseman, J. & Kirkwood, J. G. 1948 *J. Amer. Chem. Soc.* **70**, 2820.

Sandow, A. 1947 *Ann. N.Y. Acad. Sci.* **47**, 895.

Scatchard, G. 1949 *Ann. N.Y. Acad. Sci.* **51**, 660.

Scatchard, G., Scheinberg, I. H. & Armstrong, S. H. Jr. 1950 *J. Amer. Chem. Soc.* **72** (in the Press).

Schäfer, H. & Göpfert, H. 1937 *Pflüg. Arch. ges. Physiol.* **238**, 484.

Schmitt, F. O., Bear, R. S., Hall, C. E. & Jakus, M. A. 1947 *Ann. N.Y. Acad. Sci.* **47**, 665.

Schramm, G. & Weber, H. H. 1942 *Kolloidzschr.* **100**, 242.

Schulz, G. V. 1947 *Z. Naturf.* A, **2**, 348.

Snellmann, O. & Erdös, T. 1948 *Nature,* **161**, 526.

Snellmann, O. & Tenow, M. 1948 *Biochim. Biophys. Acta,* **2**, 384.

Straub, F. B. 1942 *Stud. Inst. Med. Chem. Univ. Szeged,* **2**, 3.

Straub, F. B. 1943 *Stud. Inst. Med. Chem. Univ. Szeged,* **3**, 23.

Szent-Györgyi, A. 1942 *Acta Physiol. Scand.* **9**, suppl. 25.

Szent-Györgyi, A. 1947 *Chemistry of muscular contraction.* New York: Academic Press.

Szent-Györgyi, A. *et al.* 1942–43 *Stud. fr. Inst. Med. Chem. Univ. Szeged* (ed. Szent-Györgyi). Basel & New York: Karger.

Szent-Györgyi, A. *et al.* 1946–48 *Hungar. Acta Physiol.* **1**.

Szent-Györgyi, A. 1949 *Biol. Bull.* **96**, 140.

Varga, L. 1946 *Hungar. Acta Physiol.* **1**, 1.

Weber, H. H. 1949 *Fiat Reviews,* Physiol. **3**, sect. Muscle.

Weber, H. H. 1930 *Biochem. Z.* **217**, 430.

Weber, H. H. 1939 *Naturwissenschaften,* **27**, 33.

Woehlisch, E. 1940 *Naturwissenschaften,* **28**, 305, 326.

Wolpers, C. 1944 *Dtsch. med. Wschr.* **29/30**, 495.

The stability of D-arabinose adaptation of
Bact. lactis aerogenes

By Sir Cyril Hinshelwood, F.R.S., and S. Jackson

Physical Chemistry Laboratory, University of Oxford

Bact. lactis aerogenes first subcultured into D-arabinose has a long lag phase, in one strain about 30 hr. This is shown by cultures freshly grown from a single colony, and is not significantly changed by repeated subculture in a variety of media free from D-arabinose.

According, therefore, to the theory that mutations and reverse mutations lead to an equilibrium and that the lag is the time required for the small proportion of mutants to multiply in the D-arabinose, this proportion must be very small, and the equilibrium rapidly established. Differential equations can be set up to express the rate of establishment of equilibrium starting either with normal cells (not utilizing D-arabinose) or with mutant cells. From the experimental observations an estimate can be made of the minimum rate of establishment of equilibrium in the forward direction, and the equations then can be applied to calculate the rate at which reverse mutation should occur. The calculations can be compared with the experimental results on the stability of D-arabinose adaptation during subsequent culture of the mutant cells in glucose.

Reversion does not occur in the predicted manner.

The theory would also fail to account for the generally observed influence of the length of training of bacteria on the ease of reversion.

Theories of direct adaptation do not present these difficulties (though they do not themselves make any general prediction about the rate of reversion).

Introduction

Bacteria, as is well known, adapt themselves readily to utilize new sources of carbon and nitrogen and to resist the action of toxic substances. The word adaptation may be used here without any implication about the nature of the process, which might consist either in the selection of mutant cells or in structural changes in the cell material called forth by the chemically different medium in which synthesis occurs. This is a matter upon which deep-seated differences of opinion exist* (though it may be said that the mechanism does not need to be identical in every example).

Cultures which have undergone what is called training to new conditions may lose their acquired properties (reversion), or retain them tenaciously (stable adaptation) on transfer to media in which the real or apparent stimulus is no longer present. The relations of adaptation and reversion have proved in a number of examples to conform to a well-defined pattern which is as follows. If the training is short, the reversion is often relatively easy (even though the new property had been completely acquired). If the training is more thorough the reversion is delayed, and if it is suitably prolonged very high degrees of stability may be attained (for references see Hinshelwood (1946)).

In the classical example of *Bact. coli mutabile* the lactose variants seldom revert at all (Massini 1907). With *Bact. lactis aerogenes*, which shows no special lag in lactose but requires training to optimum growth rate, the adaptation survives fifty or more

---

* For references see previous paper (Jackson & Hinshelwood 1949).

passages through glucose provided that a large number of subcultures in lactose have been made. If only a few have been made, although the mean generation time falls from an initial 100 min. to about 33, it will revert to about 80 min. after various passages through glucose (Postgate & Hinshelwood 1946).

Reversion must depend upon either (*a*) reverse mutations, (*b*) reselection of the original type of cell, (*c*) restoration of the original proportions and composition of cell material by growth in the old conditions.

Hypothesis (*a*) has received a good deal of support; and the object of the present study is to examine its applicability to the example of *Bact. lactis aerogenes* trained to use D-arabinose as the source of carbon.

BACT. LACTIS AEROGENES AND D-ARABINOSE

When first transferred from glucose to D-arabinose media the cells show a long lag, which with the strain employed was about 30 hr. After one subculture this had been practically abolished (for experimental details see Koser & Vaughan 1937; Cooke & Hinshelwood 1947; Jackson & Hinshelwood 1948). The mean generation time in the sugar is 75 min. and is steady.

We are going to examine the hypothesis that the long lag represents the time during which a minute proportion of mutants multiplies. In the experimental determination of lag a total cell number n_0 multiplies to n_t in time t, and the lag is calculated from the relation

$$L = t - \frac{1}{k_A} \ln \frac{n_t}{n_0}.$$

n_t is many times greater than n_0. If in fact the whole time is needed for the multiplication of m_0 mutants

$$t = \frac{1}{k_A} \ln \frac{n_t}{m_0},$$

n_0 being neglected in comparison with n_t. Therefore

$$k_A L = \ln \frac{n_0}{m_0}.$$

k_A is the growth rate constant in D-arabinose. The mean generation time being 75 min.

$$2 = \exp [75 k_A],$$

whence we find $\qquad m_0/n_0 = 10^{-7 \cdot 2}.$

The fraction of mutants according to the hypothesis will therefore be taken as approximately one in 10^7.

(If reverse mutations occur during growth in the D-arabinose the number, M, of the mutant cells will increase according to the equation

$$dM/dt = k_0 M - \mu' k_0 M$$
$$= k_0 (1 - \mu') M = k_A M.$$

It is, however, k_A and not k_0 which is measured in the normal determination of growth rate, and which corresponds to the 75 min. mean generation time.)

There are, according to the mutation theory, only two possible explanations for the smallness of this proportion. Either the equilibrium ratio of mutants to normal cells is small, in which case the rate of reverse mutation must be much greater than that of forward mutation; or else the culture showing the 30 hr. lag has a proportion of mutants which has not yet reached equilibrium.

If the statistical equilibrium has not yet been established then further subculture should change the proportion of the mutants and cause a progressive change in the lag in D-arabinose. Numerous subcultures through a variety of media cause no systematic increase or decrease in the lag, as is shown by the experimental results in table 1. We must therefore conclude that the mutations and back mutations are, according to the terms of the theory, already in statistical equilibrium in the standard culture.

TABLE 1. LAG IN D-ARABINOSE MEDIUM TESTED AFTER SERIAL
SUBCULTURE IN VARIOUS MEDIA

lag (hr.)

number of subcultures	glycerol (1)	glycerol (2)	lactose	maltose	cellobiose	sucrose	L-arabinose	mannose	xylose	galactose	glucose	acetate	fumarate	succinate
4	31	24	28	36	30	29	30	29	30	29	21	28	22	22
6	38	—	34	35	33	34	38	30	38	30				
8	38	—	36	36	36	38	36	36	36	36				
9	—	36	—	—	—	—	—	—	—	—	19	33	33	24
10	35	—	35	35	35	38	35	35	35	35				
16	—	38	—	—	—	—	—	—	—	—	31	28	28	30
20	33	—	33	33	38	33	33	33	33	33				
24	—	37	—	—	—	—	—	—	—	—	22	34	30	27
29	35	—	31	29	34	32	31	32	33	29				
31	—	38	—	—	—	—	—	—	—	—	24	36	33	29
35	38	—	30	35	—	38	38	33	32	33				
41	30	—	30	30	—	30	30	30	35	35				
45	—	27	—	—	—	—	—	—	—	—	34	24	—	24
51	—	24	—	—	—	—	—	—	—	—	27	24	—	25
52	38	—	31	22	—	33	24	25	24	24	—	—		
58	—	24	—	—	—	—	—	—	—	—	34	28	—	28
69	—	25	—	—	—	—	—	—	—	—	25	35	—	28
74	—	22	—	—	—	—	—	—	—	—	29	35	—	—

Since in this equilibrium the proportion of mutants is very small, one of two conclusions would follow: either the rate of forward mutation is very small, or the rate of back mutation is very great. There seemed to be a possibility that the rate of forward mutation might not be very great, since in a previous series of experiments (Jackson & Hinshelwood 1948) with another strain showing an initial D-arabinose lag of 44 hr. considerable increases were sometimes observed after subculture on agar and passage through bouillon, and in any case strains of *Bact. lactis aerogenes* have been met with lags from 1 to 9 days. These differences may quite well depend upon the enzymatic make-up of the cells, and the influence of the cultural history may act through this factor. It might, however, act by producing cultures at varying degrees of removal from equilibrium.

It was necessary, therefore, to form an idea of the rate of establishment of the equilibrium with the culture at present under study. This was done in the following way. The culture was plated on solid agar medium and individual colonies, each of which had sprung from a single cell, or at most one or two adhering cells, were selected, and subcultured once in the standard glucose-ammonium sulphate-phosphate medium. Since, by hypothesis, the original culture contained one mutant in 10^7, the chance that any of these colonies contained mutants was very small. Yet the procedure described yielded cultures all of which showed the characteristic behaviour with D-arabinose (table 2). The conclusion is that the equilibrium is established at least fairly rapidly.

TABLE 2. LAGS OF CULTURES GROWN FROM SINGLE COLONIES

colony	lag in glucose (hr.)	lag in D-arabinose (hr.)
1	3	38
2	2	32
3	2·5	29
4	0	32
5	1	32
6	3	28

We may take it that the procedure described has yielded cultures containing not less than about half the equilibrium proportion of mutants (or the lags would not be normal). The single cells multiplied first to form a colony, the whole of which was transferred to the glucose medium, in which further growth occurred until 10^{10} cells had been produced from the original one. We may conclude therefore that with the strain under investigation the mutation rate is such that in the course of a number of divisions sufficient to give a cell multiplication factor of 10^{10} the equilibrium proportion of mutants is at least half established. This gives a lower limit for the rate of the forward mutation.

EQUATIONS FOR MUTATION AND REVERSE MUTATION

It is now necessary to establish the equations for forward and back mutation, from which the consequences of this result will be seen.

Let N = the number of normal cells and M = the number of mutants. Let k be the growth rate constant in the standard glucose medium, which is in fact the same for the normal and for the 'trained' cells. Let the chance that a given new cell formed be a mutant be μ and the chance that when a mutant cell divides, a daughter cell should suffer a back mutation be μ', then we have the equations

$$dM/dt = \mu k N + kM - \mu' kM, \quad dN/dt = kN + \mu' kM - \mu kN.$$

Let $R = M/N$, so that

$$\frac{dR}{dt} = \frac{1}{N}\frac{dM}{dt} - \frac{M}{N^2}\frac{dN}{dt} \quad \text{and} \quad \frac{dR}{dt} = \mu k \left\{ 1 - \left(\frac{\mu'}{\mu} - 1 \right) R - \frac{\mu'}{\mu} R^2 \right\}. \tag{1}$$

At equilibrium $dR/dt = 0$, and

$$1 - (\beta - 1)R_{eq.} - \beta R^2_{eq.} = 0, \qquad (2)$$

where $\beta = \mu'/\mu$.

The solution of this equation gives β. $R_{eq.}$ is small. In these circumstances the equation becomes

$$1 - (\beta - 1)R_{eq.} = 0,$$

whence $R_{eq.} = 1/\beta$, the unit being negligible compared with β.

We have found that at equilibrium $M/(M+N) \sim 10^{-7}$ and therefore

$$M/N = R_{eq.} \sim 10^{-7},$$

since when R is small M/N and $M/(M+N)$ are indistinguishable.

Since

$$R_{eq.} \sim 10^{-7}, \quad \beta \sim 10^7.$$

ESTABLISHMENT OF A MUTANT POPULATION

If we start with a pure strain and follow the establishment of the equilibrium population of mutants, equation (1) may be used in the form

$$\frac{dR}{dt} = \mu k(1 - \beta R), \qquad (3)$$

since R^2 is negligible in comparison with R and 1 with β. Integration of (3) with the condition that $R = 0$ when $t = 0$ gives

$$R = \frac{1}{\beta}(1 - e^{-\beta \mu k t}). \qquad (4)$$

We have already concluded that the time required for R to assume a value of the order of magnitude of $\frac{1}{2}R_{eq.}$ is not greater than that during which the cells multiply by a factor of 10^{10}. For this

$$kt = \ln 10^{10} = 2 \cdot 3 \log_{10} 10^{10} = 23 \text{ approx.}$$

Thus in equation (4) we have

$$\frac{0 \cdot 5}{\beta} = \frac{1}{\beta}(1 - e^{-23\beta\mu}),$$

and from this we conclude that $\beta\mu$ must be greater than $0 \cdot 03$.

REVERSION OF A MUTANT POPULATION

If a pure or nearly pure strain of the mutant cells is prepared by selection, in the present example, that is, by growth in a D-arabinose medium, in which, in terms of the hypothesis under examination, the normal cells do not grow, and these are then cultured in the normal medium without the D-arabinose, the course of the population changes will be governed by the equation

$$\frac{dR}{dt} = \mu k\{1 - \beta R - \beta R^2\},$$

with the condition that $R = \infty$ when $t = 0$. The solution is

$$-\mu kt = \frac{1}{(\beta^2 + 4\beta)^{\frac{1}{2}}} \ln \left\{ \frac{2R\beta + \beta - (\beta^2 + 4\beta)^{\frac{1}{2}}}{2R\beta + \beta + (\beta^2 + 4\beta)^{\frac{1}{2}}} \right\}.$$

With a high degree of approximation the solution is

$$e^{-\beta \mu kt} = \frac{(R - 1/\beta)}{R + 1},$$

whence

$$R = \frac{e^{-\beta \mu kt} + 1/\beta}{1 - e^{-\beta \mu kt}}. \tag{5}$$

Equation (5) gives $R = 1/\beta$ for $t = \infty$ in accordance with the approximate equilibrium condition already derived.

REVERSION OF D-ARABINOSE MUTANTS

A strain which has been grown several times in D-arabinose and which shows a lag of nearly zero will, by hypothesis, have a value of $M/(M + N)$ nearly equal to unity, i.e. a value of $R \to \infty$. Its reversion will be governed by equation (5). According to this equation the question whether the non-mutant cells have been completely eliminated or not is of no consequence, since the behaviour of a culture with $R = 1 \cdot 0$ would in fact differ little from that with $R = \infty$.

When R has fallen as a result of back mutations to a small value the lag of the culture will, on renewed testing in D-arabinose, be found to be L, where

$$\ln (1/r) = k_A L,$$

where $r = M/(M + N)$ is the multiplication ratio by which the remaining mutants must increase before their number attains the value used in the calculation of L from the total inoculum size. $R = M/N$, so that when r is small compared with unity

$$r = R, \quad \ln (1/R) = k_A L. \tag{6}$$

Remembering that $\ln 2 = 75 k_A$ we have

$$\log_{10}(1/R) = (L \log_{10} 2)/75. \tag{7}$$

L is here expressed in minutes but will be converted into hours. From the previous section $\beta \mu > 0 \cdot 03$.

From (7) and (5), L, the lag developed, can be calculated in terms of kt, where t is the period for which logarithmic growth (with reversion) has been continued in glucose. kt is conveniently expressed in terms of the number of subcultures during which the reversion has been given the chance to occur. At each subculture there is under the experimental conditions employed a multiplication ratio of 10^2 to 10^3. The lower value is more conservative. Therefore in S subcultures the multiplication ratio will be 100^S. This is related to kt by the equation

$$\ln 100^S = kt. \tag{8}$$

Thus from (5), (7) and (8), L can be calculated in terms of S.
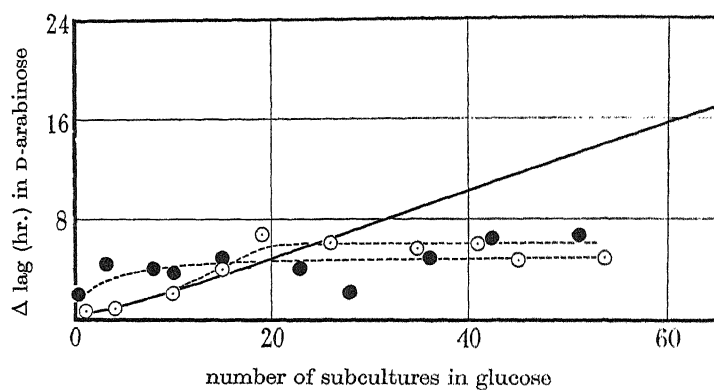
FIGURE 1. Stability of D-arabinose adaptation. Continuous line, calculated
reversion rate. Dotted lines, observed results for 2 cultures.

TABLE 3. EFFECT ON D-ARABINOSE ADAPTATION OF CULTURE IN GLUCOSE

Cells grown for n subcultures in glucose medium after x subcultures in D-arabinose medium.

	n	lag in D-arabinose (hr.)	lag in glucose (hr.)	D-arabinose lag minus glucose lag (hr.)
$x = 1$	1	3	2·5	0·5
	4	3	2·5	0·5
	10	4	2	2
	15	6	1·5	4·5
	19	8	1	7
	26	7	1	6
	35	8	2·5	5·5
	41	6	0	6
	45	5·5	1	4·5
	53	5·5	1	4·5
$x = 13$	0	4	1·5	2·5
	2	9	1·5	7·5
	6	7	1	6
	12	9	2	7
	16	14·5	1·5	13
	20	15·5	1	14·5
	24	5	2	3
	28	7	1	6
	35	6	1·5	4·5
	47	9	1·5	7·5
$x = 30$	0	3	1	2
	3	5·5	1	4·5
	8	5	1	4
	10	5	1·5	3·5
	15	6	1	5
	23	6	2	4
	28	5	3	2
	36	6	1	5
	42	8	1·5	6·5
	51	7·5	1	6·5

The values of the lag which should be observed after various numbers of sub-cultures for the case where $\beta\mu = 0{\cdot}03$ are shown in figure 1. The continuous line gives the minimum rate of reversion which should be observed, since a minimum value for $\beta\mu$ has been used.

The dotted lines shown in the figure represent the experimental results from table 3. These were obtained with cultures first trained to D-arabinose, then passed repeatedly through glucose and tested at intervals.

It is evident that the calculations do not correspond at all with the observations.

In particular, the general form of the curve, as distinct from the numerical value of the reversion rate, is incorrectly predicted.

Although it is not evident in the present example, there is often a very marked dependence of the stability of the trained culture on the number of times it has been passed through the training medium. This would be quite unaccounted for by any theory on the lines of that explored in the foregoing calculations.

If we adopt the view that training involves definite enzymatic changes in the cell and the development of new structural patterns in the cell material, then the completeness with which all traces of the old pattern had been removed by repeated growth in the new medium could well be imagined to determine whether or not the original pattern is easily restored on retransfer to the former medium. This matter will not be discussed in detail here, since the experiments were not in any way designed to throw positive light upon this form of theory.

An explanation of the non-reversion could be provided in terms of the *ad hoc* assumption that in presence of arabinose-adapted cells, any back mutants (that is, detrained cells) were unable to multiply at a normal rate. Since, in fact, the mean generation times in glucose of trained and untrained cells are identical, this assumption is extremely arbitrary, and devoid in the present case of independent support. An assumption of this type, moreover, would be of little help in understanding the general phenomena of unstable adaptation, or of reversion occurring after variable delays.

REFERENCES

Cooke, E. A. & Hinshelwood, C. N. 1947 *Trans. Faraday Soc.* **43**, 733.
Hinshelwood, C. N. 1946 *Chemical kinetics of the bacterial cell.* Oxford: Clarendon Press.
Jackson, S. & Hinshelwood, C. N. 1948 *Trans. Faraday Soc.* **44**, 568.
Jackson, S. & Hinshelwood, Sir Cyril 1949 *Proc. Roy. Soc.* B, **136**, 562.
Koser, S. A. & Vaughan, E. F. 1937 *J. Bact.* **33**, 587.
Massini, R. 1907 *Arch. Hyg., Berl.*, **61**, 250.
Postgate, J. R. & Hinshelwood, C. N. 1946 *Trans. Faraday Soc.* **42**, 45.

The initiation of nerve impulses by mesenteric Pacinian corpuscles

BY J. A. B. GRAY AND J. L. MALCOLM

National Institute for Medical Research, Hampstead, London

A preparation of a single Pacinian corpuscle in the cat's mesentery has been used to study the initiation of nerve impulses in sensory endings.

The minimum movement of a mechanical stimulator required to excite a single corpuscle has been found to be 0.5μ in 100 μsec.

It has been difficult to produce repetitive discharges with rectangular pulses of long duration, either mechanical or of constant current.

The latency between a mechanical stimulus and the initiation of an impulse has a value around 1·5 msec. for threshold stimuli, and this decreases to a minimum value around 0·5 msec. as the stimulus is increased; it is altered only slightly, if at all, by changes in the duration of the maintained displacement of the mechanical stimulator.

Subthreshold mechanical stimuli have been shown to facilitate stimulation by electrical test shocks.

The return of excitability at the ending is independent of the nature of the conditioning stimulus and varies but little with the nature of the test shock. The value of the latency at threshold is unaffected by the relatively refractory state.

The relations of these results to various hypotheses are discussed, and it is suggested that these results can all be accounted for in terms of the known properties of axons.

INTRODUCTION

The pattern of the discharge from sensory receptors has been studied by a number of workers since the investigations of Adrian (1926). The mechanism determining this pattern in single mammalian endings was considered by Adrian & Umrath (1929) using the Pacinian corpuscles lying beside the tendons in the cat's foot, and later by Matthews (1933) using the stretch receptors in cat's muscle. The elucidation of the mechanism of the receptor was taken a step further by Granit and his collaborators, who suggested, as a result of experiments with retinal units, that the ending produced a 'generator potential' and that the frequency pattern in the nerve was determined by the properties of the axon, a view supported by their investigations of the repetitive properties of the nerves when stimulated by constant currents (Bernhard, Granit & Skoglund 1942; Skoglund 1942; Granit 1947). This conception is essentially the same as that put forward by Adrian (1930) when discussing the repetitive discharge from a cut end of peripheral nerve.

Further investigations of the physiological initiation of nerve impulses at sensory receptors seemed to demand a preparation in which a single sensory receptor could be stimulated and recorded from at will. Such a preparation would allow the application to sensory receptors of methods used to study the electrical excitation of nerve. The Pacinian corpuscle in the cat's mesentery seemed to provide a suitable preparation (Gray 1947).

The Pacinian corpuscle has long been the object of study by anatomists and physiologists. The first description appears to have been made by Johannes Gottlieb Lehmann, who included it in a thesis for his Doctorate read on 11 November 1741, before Abraham Vater, F.R.S., Professor of Anatomy and Botany at Witemburg. This short description of the gross anatomy seems to have been forgotten, and Pacini published his preliminary description in 1835, followed by a full account of their structure in 1840. Their physiological significance was then the subject of much speculation. Andral (1837) described them as 'ganglia of tact'. Pacini (1840) noted their apparent structural resemblance to the electrical organ of the torpedo, and made the tentative suggestion that they might be receptors of animal magnetism. Henle & Kölliker (1844) confirmed and enlarged Pacini's histological observations, but disagreed with his theory of animal magnetism. They were, however, struck with their resemblance to the electrical organ and attempted unsuccessfully to record electrical potentials from them. Todd & Bowman (1845) presumed that they subserved the sensation of touch because of the large numbers found in the hands and feet. This question of function, however, was not investigated experimentally until Adrian & Umrath (1929) recorded the action potentials in nerves from Pacinian corpuscles when pressure was applied to them with a glass rod. Their function in the mesentery was considered by Gammon & Bronk (1935) and by Gernandt & Zotterman (1946).

The results reported here are derived from experiments carried out on single Pacinian corpuscles in the hope of obtaining more information on the fundamental mechanism of sensory receptors. Our results can be explained in terms of the known properties of axons, and there is no need, in the case of Pacinian corpuscles, to postulate the existence of a receptor cell or any other mechanism intermediate between the stimulus and the axon. A preliminary report has already appeared (Gray & Malcolm 1949).

METHODS

General

The experiments were made on cats anaesthetized with 80 to 100 mg. of chloralose per kg. body weight; a few were performed on cats decerebrated under ether anaesthesia.

The abdomen was opened and the wound margins attached to a ring (*a*, in figure 1), which was clamped firmly to the operating table. A stage of thick brass, covered with cork (*c*, in figure 1), was held in the open abdomen by means of a rod (*d*, in figure 1) mounted on a heavy stand which had a fine vertical adjustment. The selected portion of the mesentery was laid on the stage and pinned in position. As soon as this had been done the abdomen was filled with warm liquid paraffin. The vertical adjustment of the stage was used to raise and lower the preparation through the paraffin surface during the course of the dissection.

Dissection

Pacinian corpuscles are distributed around the mesenteric vessels mainly in the jejunal region and occasional single ones are found in the clear mesentery (*a*, in figure 2). It is these single corpuscles which give suitable preparations; their nerves

usually accompany fine vessels to the mesenteric arch where they join the mesenteric nerve plexus.

A corpuscle was selected and the surrounding mesentery was pinned out on the stage. The dissection was then carried out under paraffin with the aid of a binocular
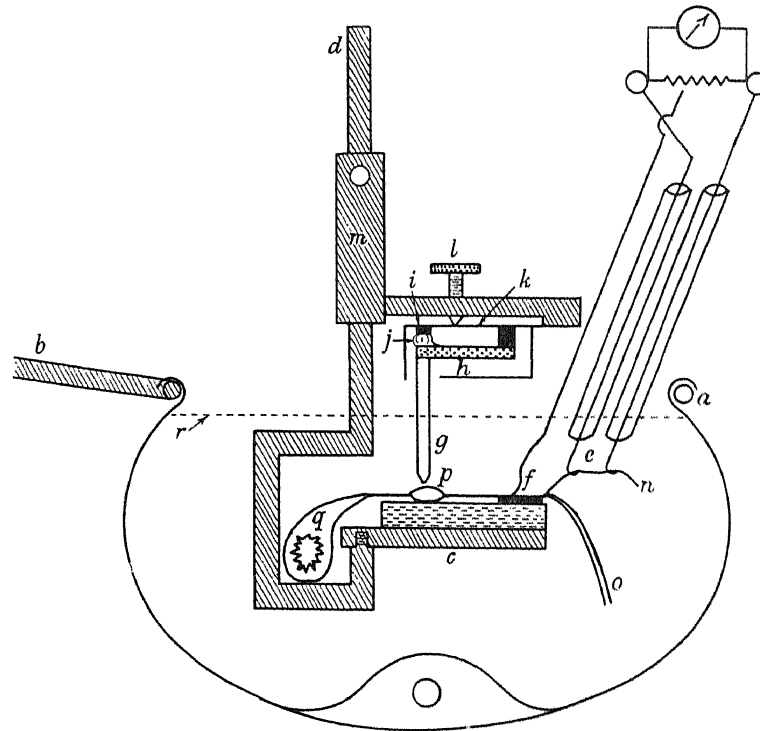
FIGURE 1. Transverse section of the abdomen to show the arrangement of the preparation. a, supporting ring; b, insulated supporting rod; c, brass stage, covered with cork; d, brass support; e, electrodes; f, platinum plate; g, glass stylus; h, Rochelle-salt crystal; i, Perspex platform; j, damping oil; k, phosphor-bronze strip; l, fine-adjustment screw; m, universal clamp; n, mesenteric nerve; o, mesenteric root; p, Pacinian corpuscle; q, intestine; r, paraffin surface.
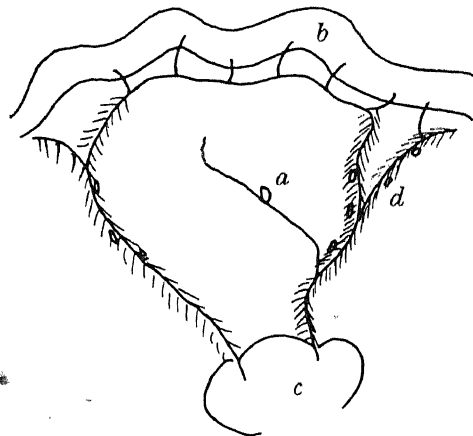
FIGURE 2. Distribution of Pacinian corpuscles in mesentery. a, isolated corpuscle; b, intestine; c, mesenteric root; d, mesenteric arch with Pacinian corpuscles.

dissecting microscope of either 12 × or 25 × magnification. The first step was to find the part of the mesenteric nerve in which the fibre from the corpuscle ran. The corpuscle and the fine vessels connecting it to the mesenteric arch were then dissected as free from mesentery as possible, except at the peripheral end of the corpuscle where a tag of mesentery was left. A small black ebonite plate was sometimes pinned to the stage under the corpuscle to facilitate its dissection. After the dissection this plate was removed and replaced by the electrode illustrated in figure 3 and described below.

Electrodes

The electrode assembly on which the corpuscle was mounted (figure 3) consisted of an 8 × 4 × 1 mm. Perspex plate, which was pinned to the cork stage, and down its centre was cut a groove in which the corpuscle could lie. Into the peripheral end of this groove a platinum wire was so fixed that its central end formed a small vertical pin (*a*, figure 3) which held the corpuscle by transfixing the tag of mesentery. A second electrode (*b*, in figure 3) was arranged to lie under the central pole of the corpuscle; it was also a platinum wire cemented into a transverse groove in the block. A platinum hook electrode was sometimes used at point *c* (figure 3).
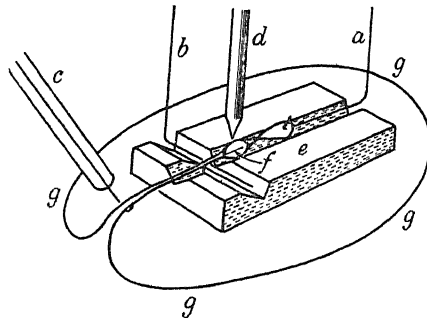
FIGURE 3. Holder for dissected Pacinian corpuscle. *a*, *b* and *c*, electrodes; *d*, glass stylus; *e*, Perspex block; *f*, Pacinian corpuscle; *g*, mesentery.

The mesenteric nerve was laid on two platinum hooks (*e*, in figure 1). The reference point of the recording system was a plate covered with platinum foil pinned to the cork under the mesenteric arch (*f*, in figure 1).

Recording and stimulating circuits

Action potentials were recorded either between electrode *e* (figure 1) or electrodes *b* and *c* (figure 3). The electrodes were connected to the grids of a pair of cathode followers, the reference point being connected to plate *f* (figure 1). The potentials were amplified and recorded on one beam of a double-beam cathode-ray oscillograph.

Two types of electrical stimuli were used, either unidirectional pulses of 30μ sec. duration, or rectangular pulses of current from a circuit with an output impedance of at least 5 MΩ. In each case a potential proportional to the voltage or to the current of the stimulus was applied to the second beam of the oscillograph. These deflexions were calibrated and used to determine stimulus strengths and intervals.

Mechanical stimulator

Mechanical stimuli were transmitted to the corpuscle through the glass stylus (*g*, in figure 1) (*d*, in figure 3), from a Rochelle-salt crystal (*h*, in figure 1). This crystal (type Rothermel 4D41), 15 mm. square, was mounted in a brass case on three of its corners. The remaining free corner was attached to the upper end of the glass stylus. A potential difference, up to 50 V, from a source having an output impedance of 1000 Ω, applied across the crystal so bent it that the free corner had a maximum vertical excursion of about 25 μ. In these experiments only rectangular potential changes were employed.

The movement of the tip of the stylus was recorded under two conditions. First it was photographed through a $\frac{2}{3}$ in. microscope objective on to a falling plate camera; under these conditions the stimulator was unloaded and the movement was maximal. Secondly, the stylus was placed on another piezo-electric crystal, and the potential generated by it during operation of the stimulator was amplified and recorded on the oscillograph; under these conditions the movement of the stylus was small.
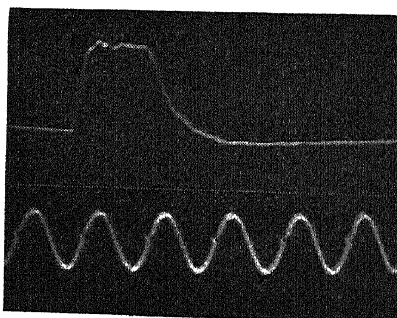
FIGURE 4. Record of the potential developed in a recording crystal during the application of a mechanical pulse from the stimulator. Time, 3 kc./sec.

Both methods of recording the stimulus movement showed that damping was required. This was obtained by placing a platform (*i*, in figure 1) under the free corner of the crystal, so as to leave a gap of approximately 0·5 mm., which was filled with oil of a suitable viscosity (*j*, in figure 1).

The results provided by the optical method showed that the damping was adequate, but the method was not sensitive enough for calibration or detailed consideration of the wave form. Figure 4 is an oscillograph record of the potential developed in the recording crystal during the application of a 400 μsec. 'rectangular' pulse from the stimulator. It will be seen that the rising time is about 80 μsec., and that it is followed by a few highly damped oscillations. It is well known that the movement of these crystals is linearly related to the applied voltage; it was found that this relationship still held in the complete stimulator.

The crystal and its case were mounted on a phosphor-bronze strip which was firmly fixed at one end (*k*, in figure 1). A screw (*l*, in figure 1) working against the free end allowed fine adjustment of the vertical position. All other movements were provided by the universal clamp (*m*, in figure 1) which held the stimulator in

position on the rod supporting the stage. In this way stage and stimulator were rigidly clamped together. The recording of the stimulus form discussed above was carried out on this mounting, the recording crystal being fixed to the stage.

The stylus tip was rounded and of the same order of diameter as the corpuscle itself. It was always applied to the side of the corpuscle (figure 3).

RESULTS

Threshold

A rectangular mechanical pulse applied to a corpuscle sets up action potentials in the axon at the 'on' and the 'off' of the pulse. When the pulse is of long duration, the threshold at the 'on' is usually lower than the threshold at the 'off'. With short durations the two can interact; the 'off' stimulus can occur in the refractory period of the action potential resulting from the 'on' stimulus, and sometimes a subthreshold 'on' can summate with a subthreshold 'off'.

We have estimated the minimum extent and velocity of movement of the stylus which can excite. A typical value is 0.5μ in 80 μsec., which equals a velocity of 0.006 m./sec.

Response to repetitive stimuli

The behaviour of a corpuscle stimulated repetitively with short mechanical pulses (< 500 μsec.) was investigated. It was found that the minimum interval between action potentials for any given stimulus strength agreed with the values expected from direct measurement of the refractory period (1 to 2 msec.). In one preparation a stimulus frequency of 1000 c./sec. was successfully followed for many seconds. When the interval between stimuli was reduced to less than the refractory period, the corpuscle responded only to every second, third or fourth stimulus.

Repetition during stimuli of long duration

A few observations were made on the repetition of action potentials during long mechanical and electrical stimuli. The scope of these experiments was limited to the use of rectangular pulses and by the maximum movement of the mechanical stimulator. The results are summarized in table 1. It can be seen that of the four preparations tested with mechanical stimuli over a range of amplitudes of stylus movement up to 30 times the threshold value, only one gave more than one action potential.

In the experiments in which electrical stimuli were used, constant currents were applied to the axon at various points on its course, usually within a few millimetres of the corpuscle. The cathode was situated nearer the recording electrodes, and controls were carried out with the animal freed from earth except for the path through the amplifier; this control was necessary, since the anode of the stimulator was earthed and anodal block could limit the repetitive discharge. The limited number of action potentials and the size of the stimulus required to produce a second action potential (table 1) contrast with the observations of Blair & Erlanger (1938) and Skoglund (1942) on other sensory fibres. The stimulus strength given in column 6 of table 1 is that which gave the maximum number of impulses;

any increase in stimulus above this optimum caused a decrease in the number of action potentials, presumably through cathodal block. With mechanical stimulation, the maximum number of impulses was limited by the size of the largest available stimulus (see methods).

TABLE 1. REPETITION PRODUCED BY LONG ELECTRICAL AND
MECHANICAL STIMULI

1 experiment no.	2 minimum stimulus strength to produce 2 action potentials	3 time between 1st and 2nd action potentials (msec.)	4 minimum stimulus strength to produce 3 action potentials	5 time between 2nd and 3rd action potentials (msec.)	6 stimulus strength for max. no. of action potentials	7 maximum no. of action potentials
			electrical stimulation			
1	6·75	2·4	10·5	2·5	10·5	3
2a	2·2	—	2·65	—	6	5
2b	1	1·69	2	2·65	7·2	20
3	>3·5	2·04	>3·5	2·75	>3·5	3
7	—	—	—	—	1 to 30	1
			mechanical stimulation			
4	3·7	2·4	—	—	>30	5
5	—	—	—	—	1 to 30	1
6	—	—	—	—	1 to 35	1
7	—	—	—	—	1 to 12	1

Stimulus strengths are given as multiples of threshold.

The difference between our results and those of Skoglund (1942) might have been due to differences in technique; we have therefore repeated one of his experiments. A cat's popliteal nerve was stimulated with rectangular current pulses, using the technique we employed with the corpuscles, and action potentials were recorded in the dorsal root. We obtained results similar to those described by Skoglund.

Latency between stimulus and action potential

Figures 5 and 6 illustrate the result of an experiment in which the latency between the stimulus and the initiation of the impulse was recorded for varying stimulus strengths.

The upper curve of figure 6 is drawn through two sets of points which were obtained by stimulating a corpuscle with rectangular mechanical pulses of 170 μsec. ($\times$), and infinite duration ($\circ$). The ordinate of this curve is the time between the onset of the stimulus and the first change of potential recorded with an electrode on the corpuscle (cf. figure 5).

In the same experiment measurements were made of the latency after a 30 μsec. electrical pulse applied to the corpuscle, by recording the action potentials from the mesenteric nerve. The times so measured were corrected by subtracting the conduction time to give the points plotted in the bottom curve of figure 6.

The third curve plotted in figure 6 was derived from the results of two other

a   b
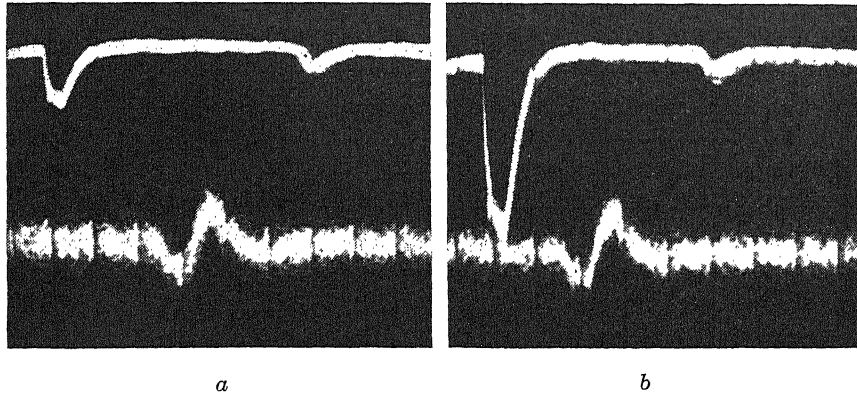
FIGURE 5. Records comprising many superimposed traces, from an experiment for measuring latency. Top trace, first deflexion, proportional to mechanical stimulus strength; second deflexion, action potential at the mesenteric nerve. Bottom trace, action potential recorded at the corpuscle. Modulation, 3 kc./sec. *a*, stimulus strength 2·7 × threshold; *b*, stimulus strength 8·7 × threshold.
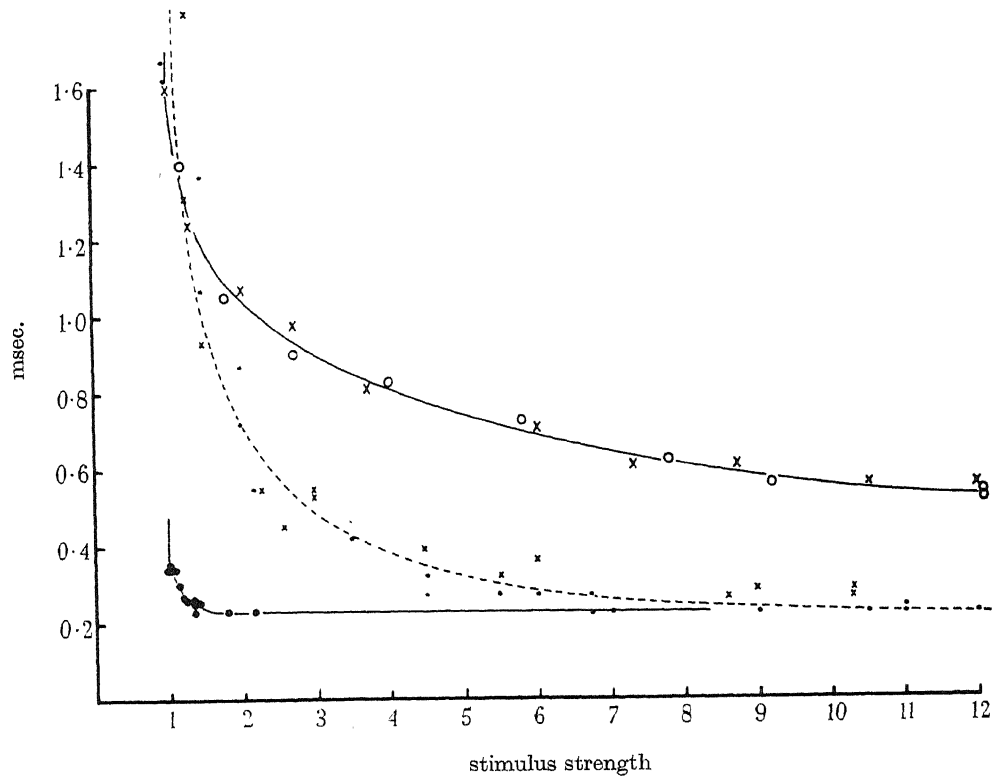
FIGURE 6. Latency curves. Abscissa: stimulus strengths in multiples of threshold. Ordinate: latency between stimulus and action potential, msec. × ———— × , mechanical stimulus, 170 μsec. duration; O ———— O, mechanical stimulus, infinite duration; ●————●, electrical stimulus, 30 μsec. duration; ●— — — —●, × — — — —× , constant current stimulus, infinite duration, from two other experiments.

experiments in which constant current pulses of infinite duration were applied between electrodes a and c (figure 3), the latter being the cathode. In the same experiments the minimum latencies were also determined with 30 μsec. pulses and were found to be identical with those produced by the pulses of infinite duration. In order that these results might be compared with those of the experiment described above, each set of figures has been corrected by the subtraction of a time such that their minimum latencies coincided with that of the bottom curve of figure 6.

It can be seen from figure 6 that in this experiment, the duration of a mechanical stimulus had no effect on the latency, whereas the difference between the effect of the two durations of electrical stimulus is striking. Such independence of duration with mechanical stimuli has not been seen in all experiments, and the results have been complicated by oscillations or overdamping of the stimulator. The shape of these curves was consistent throughout our experiments, and those shown can be regarded as typical. Table 2 gives values for the latency at threshold and the minimum latency, that is, the value obtained with a stimulus of infinite strength. Column 6 of this table has been included to give some indication of the slopes of the curves; the figure given is the strength of stimulus required to give a latency half-way between the threshold and minimum values. In two experiments (nos. 10b and 14) the latency was measured as described above. In the remaining experiments the action potentials were recorded at the mesenteric nerve, as this was technically

TABLE 2. VALUES OF LATENCY BETWEEN STIMULUS AND THE
INITIATION OF THE IMPULSE

1 exp. no.	2 latency at threshold (msec.)	3 time from which latency measured	4 method	5 minimum latency (msec.)	6 stimulus strength for a latency equal to the mean of cols. 2 and 5
A. Mechanical stimulation					
8	0·9	m.e.l.	r	—	—
9	0·8	m.e.l.	r	—	—
7	0·75	m.e.l.	d	0·39	2·7
10a	1·5	m.e.l.	r	—	—
10b	1·6	zero	d	0·48	2·0
11	0·8	m.e.l.	r	—	—
12a	1·3	m.e.l.	r	—	—
12b	1·5	m.e.l.	d	0·7	1·8
13	1·4	m.e.l.	d	0·55	1·2
14	1·7	zero	d	0·52	1·8
15a	1·3	m.e.l.	r	—	—
15b	1·3	m.e.l.	d	0·35	2·0
B. Constant current stimulation					
1	1·45	m.e.l.	d	0	1·6
2	4·45	m.e.l.	d	0	1·2
—	1·45*	—	—	—	1·55

m.e.l. = minimum electrical latency.
r   = measured from refractory period estimations (see text).
d   = direct measurement.
*   = point at which curve departed significantly from threshold.

simpler. The measured value therefore included the conduction time, and to correct for it the interval between a large electrical stimulus and its action potential was subtracted from the values obtained; the difference thus being the latency less the minimum electrical latency. The value of the minimum electrical latency is known to be small (Erlanger & Gasser 1937). Values obtained, using this as a base-line, are marked m.e.l. in table 2 column 3. In contrast with the minimum electrical latency, however, the minimum values after a mechanical stimulus are always appreciable (table 2 column 5).

Facilitation

The 'excitability' changes after subthreshold mechanical stimuli were investigated by measuring the minimum strength of a test stimulus required to initiate an impulse. Electrical stimuli of 30 μsec. duration were used as test shocks at various time intervals after the beginning of the conditioning stimulus. The results of one such experiment are given in figure 7.
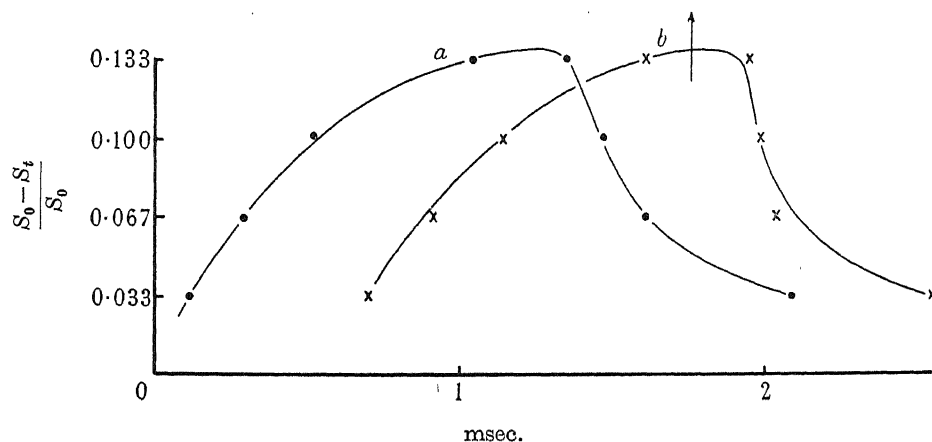
FIGURE 7. Facilitation after a subthreshold mechanical stimulus. Ordinate: $(S_0 - S_t)/S_0$, where S_0 = threshold to test shock alone and S_t = threshold at time t. Abscissa: time from beginning of conditioning stimulus to a, test shock; b, to commencement of action potential in the mesenteric nerve. The arrow marks the time at which an action potential occurred after a just threshold mechanical stimulus. Conditioning stimulus, 0·76 × threshold strength; duration, 100 μsec.

Curve a in figure 7 shows the 'excitability' increasing to a maximum at 1·2 msec. after the beginning of the conditioning stimulus. Curve b is included to illustrate the relation between this maximum and the time at which an action potential is initiated by a just threshold mechanical stimulus (arrow, figure 7). In curve b the abscissa is the time from the beginning of the stimulus to the appearance of the action potential at the recording electrodes, i.e. it is to the right of curve a by a conduction time.

It has not been possible, because of technical difficulties, to obtain a series of curves for different strengths of the conditioning stimulus, but the existence of a facilitation process has been clearly shown to follow a short conditioning pulse in three preparations. Further work on this problem is in progress.

The relationship of these curves to the time of initiation of the propagated response, and their similarity to the results of Katz (1937, 1947) and Hodgkin (1938) on subthreshold potentials in nerve, suggests that they are a reflexion, distorted by spatial spread, of subthreshold potentials at the endings.

Recovery of excitability after an action potential

General considerations

The recovery of excitability of the ending, or of the axon near the ending, has been measured with various combinations of stimuli in twelve experiments. Either short electrical or short mechanical pulses were used as first or as test stimulus, thus making four combinations possible. For convenience these will be referred to here as, for example, electrical-mechanical or mechanical-mechanical recovery curves. Unless otherwise stated the electrical stimuli were applied through electrodes a and b (figure 3), the cathode usually being peripheral; some measurements were made after an antidromic impulse initiated in the mesenteric nerve.

The first stimulus always had a strength several times the threshold. The time from it to the second stimulus was set to various values in turn, the threshold being measured at each. For the shortest interval, the test stimulus was set at its maximum strength and the time adjusted until the threshold was found. Finally, the threshold to the test stimulus was measured in the absence of the first stimulus. The results were then plotted, threshold strength as a multiple of the threshold to the test stimulus alone, against either stimulus interval or action-potential interval.

Figure 8 is designed to show the relation between the times involved in measuring the recovery curve. Line a shows the stimuli, b the action potentials at the point of stimulation and c the action potentials at the recording electrodes some distance from the corpuscle. An analysis of the times is given in the lower half, with a scale indicating their order of magnitude.

A stimulus is indicated by a dotted line at 1 msec. in figure 8, line a, to illustrate one source of error. If the absolute refractory period is assumed to be 1·2 msec., a stimulus started at 1 msec. could persist until the end of the absolute refractory period and behave like a stimulus applied at 1·2 msec. With short electrical stimuli this is probably due to the charging of tissue capacities, whereas with short mechanical stimuli it may be due to stimulation at 'off' instead of 'on'. An example of this appears in figure 9. The line through the crosses in figure 9a has been drawn vertical at that interval at which the action-potential interval reached a minimum (cf. figure 9b). The circled crosses in figure 9a to the left of this line indicate stimuli which must have persisted until the absolute refractory period terminated.

So far we have only considered the case in which the point at 1·2 msec. in figure 8 is regarded as the end of the absolute refractory period. It is possible that even in the relatively refractory period a similar error might occur if the stimulus persisted and were overtaken by a more rapid process of recovery. If this occurred, the differences between action-potential intervals and stimulus intervals would increase with decreasing stimulus intervals; there is no evidence, however, that the differences we have observed are other than random.

The differences between action-potential intervals and stimulus intervals would also increase with decreasing stimulus interval if conduction were slowed in the relatively refractory period. The random nature of these differences shows that errors due to this cause would not have been detectable in our results. This being so, the conduction times (figure 8) can be regarded as equal, hence the measured action-potential interval (figure 8, line c) can be taken as the action-potential interval at the point of stimulation (figure 8, line b). Therefore, the difference between the
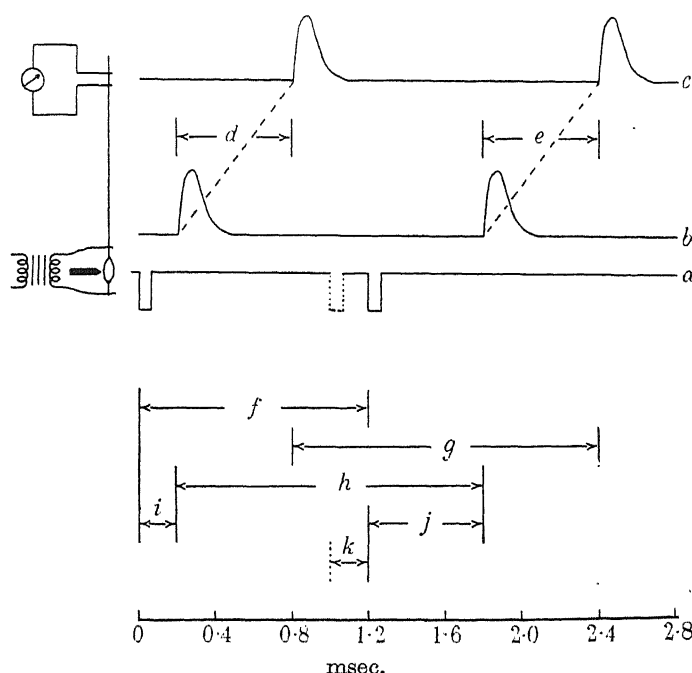
FIGURE 8. Diagram to show the times involved when measuring recovery curves and threshold latency. Line *a*, stimuli; *b*, action-potentials at point of stimulation; *c*, action-potentials some distance from the corpuscle. Interval *d*, first conduction time; *e*, second conduction time; *f*, stimulus interval; *g*, action-potential interval at distant point; *h*, action-potential interval at point of stimulation; *i*, first latency; *j*, second latency; *k*, persistence of stimulus during the absolute refractory period.

action-potential interval and stimulus interval is the difference between the latencies after the second and after the first stimulus (figure 8). It is clear that if the first stimulus is a suprathreshold electrical pulse this difference between action-potential interval and stimulus interval will be the latency following a threshold stimulus less the minimum electrical latency. Values obtained in this way are given in table 2 and marked 'r' in column 4; they are comparable with the other values which are marked m.e.l. in this table and have been discussed above. In experiments 10, 12 and 15 values have been obtained by both methods.

The stimulating electrodes in these experiments were arranged so as to stimulate as near the peripheral end of the nerve as possible. If the sites at which the electrical and mechanical stimuli initiated impulses had been different, then the differences between the action-potential intervals and stimulus intervals would have been

J. A. B. Gray and J. L. Malcolm

increased by the conduction time between these sites. Since there was good agreement between the observed differences and the directly measured latencies, this error can be ignored. Another indication that the electrical stimulus is effective as far peripheral as possible is that reversing the electrodes does not alter the recovery curves.

It is the action potential and not the stimulus which renders the nerve refractory, and it is therefore the times from the first action potential that are of interest in considering the recovery process; the action-potential interval is the most practicable
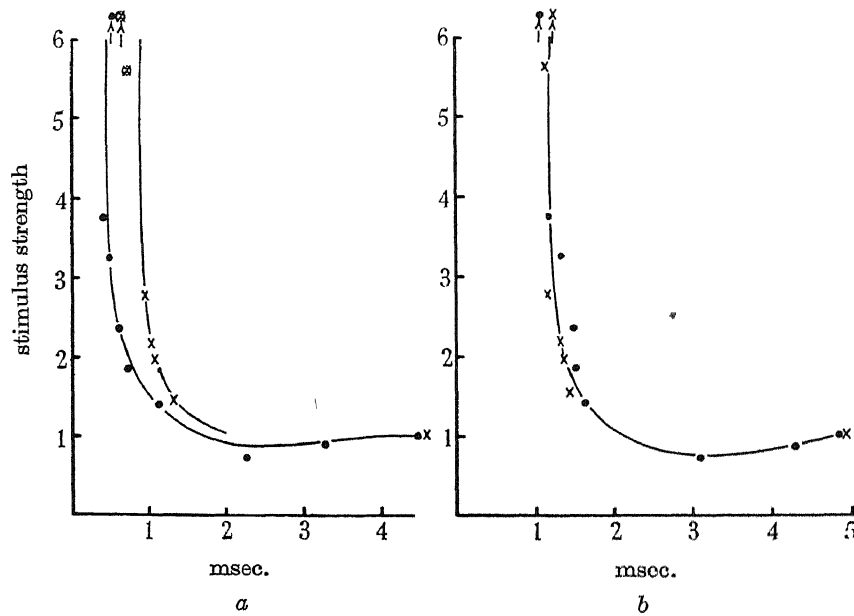
FIGURE 9. Recovery curves, using a mechanical test shock. Ordinates: stimulus strength as multiple of threshold. Abscissae: a, stimulus interval in msec.; b, action-potential interval in msec. •, electrical conditioning stimulus. ×, mechanical conditioning stimulus.

value to use. The time between the initiation of the first action potential, and the second stimulus is equally important, but demands a knowledge of the conduction time between stimulating and recording electrodes. The whole period of the second latency (figure 8), during which the second action potential is developing, is influenced by the recovery processes caused by the first. We have therefore plotted our recovery curves as bands; these lie between the curve relating action-potential interval to stimulus strength and the same curve displaced to the left by the appropriate latency (figure 10).

The nature of the first stimulus and the recovery

In six comparisons, the recovery curves were unaffected by the nature of the first stimulus. Thus, mechanical-mechanical and electrical-mechanical curves derived from the same preparation were identical (figure 9b), as also were electrical-electrical and mechanical-electrical recovery curves.

The nature of the second stimulus and the recovery

Figure 10 illustrates a typical comparison of recovery curves obtained with electrical or mechanical test shocks. The results are plotted as bands, one latency wide, and in all five experiments the band for an electrical test shock lies within the band for a mechanical one. In three experiments, the 'electrical' band lay near the
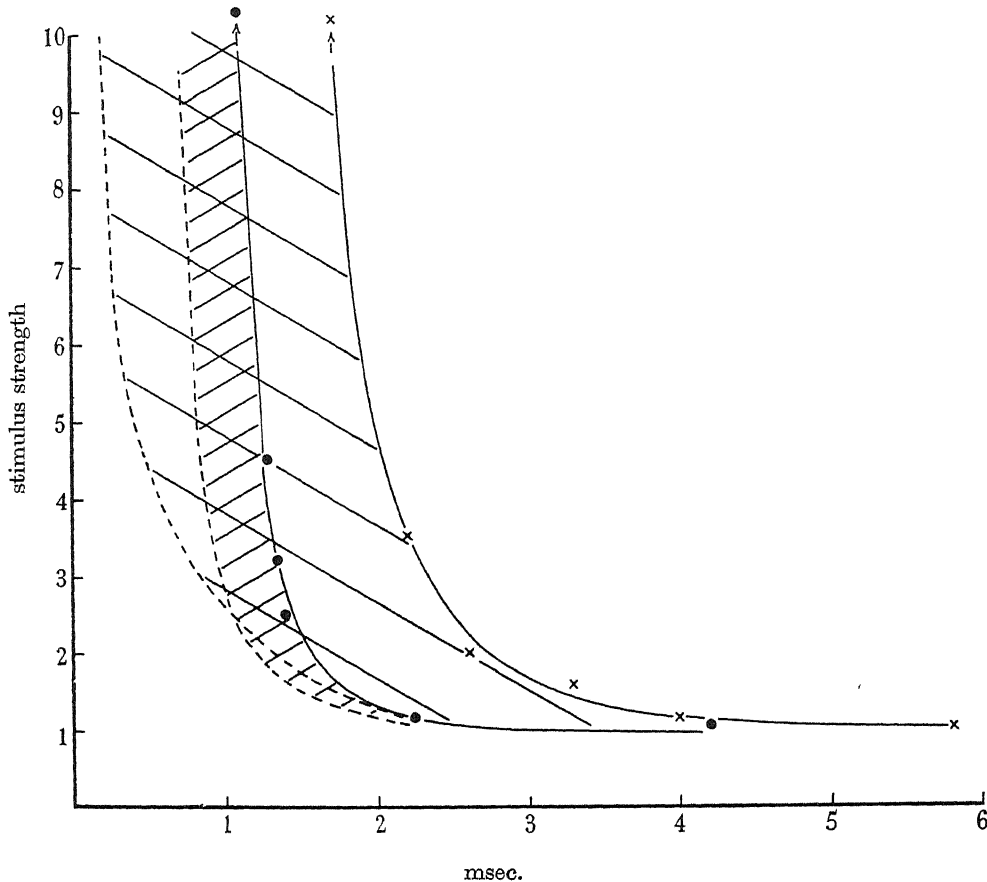
FIGURE 10. Recovery curves using an electrical conditioning stimulus. Ordinate: stimulus strength as multiple of threshold. Abscissa: full lines, action-potential interval; dotted lines, action-potential interval minus threshold latency (1·5 msec. for mechanical, 0·4 msec. for electrical). •, electrical test shock. ×, mechanical test shock.

centre of the 'mechanical' band, in one it lay along one edge, and in the remaining one along the other. The values of the absolute refractory period tested with mechanical stimuli ranged from 1·1 to 2·6 msec. with a mean of 1·9 msec. (nine experiments).

Recovery after an antidromic action potential

To eliminate the possibility of structures other than the axon being made refractory, the first action potential was set up in the mesenteric nerve and was thus conducted to the corpuscle antidromically. A smooth recovery curve was

obtained after such a conditioning action potential, showing that the main recovery process occurred in the axon, and that the stimulus applied to the axon as a result of the mechanical disturbance must have been graded.

Latency during recovery process

Figure 11 is an example of an experiment in which a mechanical test stimulus was given at intervals after the corpuscle had been made refractory by an anti-dromic impulse: the threshold of the corpuscle to the mechanical stimulus was measured, and the latencies of its responses were determined at threshold strength and at various multiples of it. One set of points was determined when the threshold
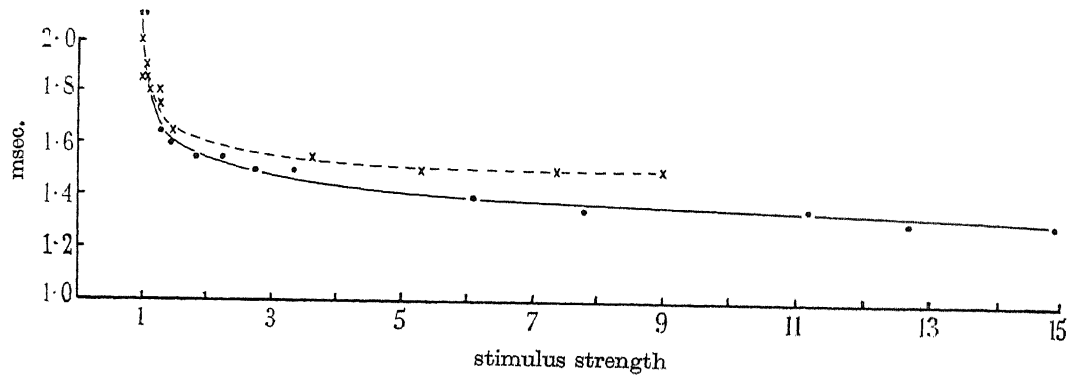
FIGURE 11. Latency after a mechanical stimulus. •, normal; ×, after the threshold had been raised to 2·1 normal by an antidromic conditioning impulse. Abscissa: stimulus strength in multiples of the threshold for each curve. Ordinate: latency (+conduction time) msec.
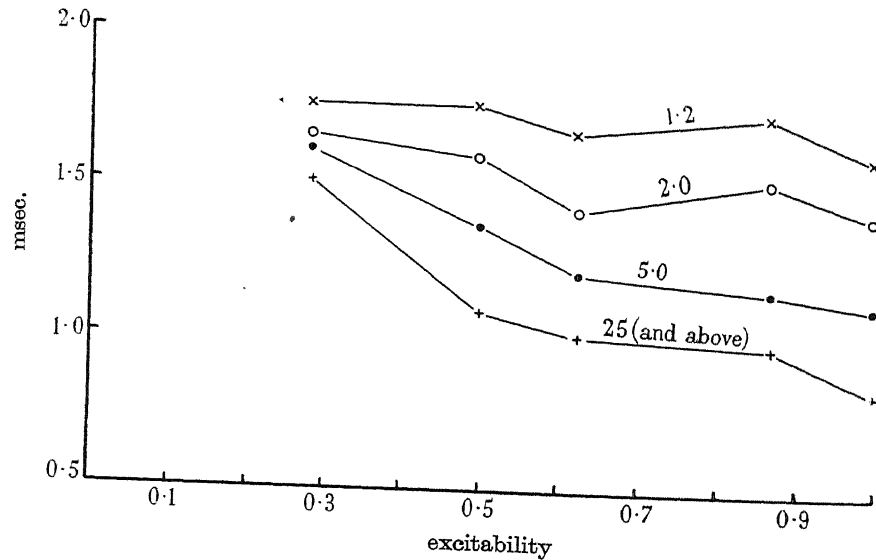
FIGURE 12. The relation of latency to excitability during the refractory period. Abscissa: excitability as the reciprocal of threshold stimulus strength. Ordinate: latency (+conduction time) in msec. Lines for different stimulus strength, given in the figure as multiples of their own threshold.

of the corpuscle was normal, i.e. outside the refractory period, and another set with the two stimuli at such an interval apart that the threshold was raised 2·1 times. It is clear from figure 11 that, when the stimulus strengths for each curve are plotted as multiples of their own threshold, the latency curves agree well for the smaller stimuli, although there is some divergence with large ones. If the absolute magnitude of the mechanical stimulus had been used as the abscissa, it is obvious that the two curves would have been widely separated.

An estimate of the latency after a threshold mechanical stimulus is given by the difference between the action-potential interval and the stimulus interval of an electrical mechanical recovery curve (see above). Such a recovery curve therefore gives a series of estimates of this latency for different stimulus intervals. As has already been stated, in none of our experiments have the latencies so measured altered with different values of stimulus interval; that is they have not been affected by the state of recovery of the axon.

A more detailed analysis of the changes in latency during the relatively refractory period is given in figure 12, in which latency is plotted against excitability for various strengths of test shock.

DISCUSSION

It is possible to draw certain conclusions from these results, and at the same time to suggest a mechanism which could explain the facts. Our finding that the latency of the response after a mechanical stimulus is little affected by the duration of the stimulus suggests that the corpuscle is more sensitive to change than to a constant deformation (cf. Katz 1949). It must be remembered that the maximum movement of the stylus used in these experiments was of the order of 25 μ, approximately 40 times threshold, and that only sharply rising pulses have been used.

It will immediately be noted that our preliminary observations on repetition differ from the results of Adrian & Umrath (1929), who obtained repetitive discharges lasting many seconds. This difference may well be due to the small size of the maximum deformation we used, although it is possible that it is due to differences between mesenteric and joint corpuscles. Such a difference is not unlikely if the nature of the repetition is decided by the axon and not by the mechanical structure of the corpuscle. Adrian & Umrath suggested that this might be so, and the idea has recently been amplified by Bernhard *et al.* (1942). Our results are not inconsistent with this, since it was difficult, in terms of threshold, to produce repetition by constant current stimulation of the axon (table 1), whereas Skoglund (1942) has shown, and we have been able to confirm, that some sensory axons from the lower limb can produce prolonged discharges with current strengths 1·5 to 3 times threshold. To answer these questions about repetitive discharge it will be necessary not only to try bigger deformations, but also to work out the accommodation of axon and corpuscle to slowly rising currents or pressures.

A smooth recovery curve is obtained when a corpuscle, rendered refractory by an antidromic action potential, is tested with a mechanical stimulus. If it is

assumed that an antidromic impulse does not pass beyond the axon, this result must mean that the stimulus applied to the axon is graded, and not 'all or none'. If this is considered in conjunction with the evidence that the nature of the conditioning stimulus makes no detectable difference to the recovery curve, it is probable that the recovery process occurs solely in the axon. The alternative demands the existence of an intermediate mechanism which can become refractory. Such an intermediate mechanism would probably be excited by the supra-threshold electrical conditioning stimulus. As there is a difference in latency after mechanical and electrical conditioning stimuli of the order of 0·8 msec., an intermediate mechanism would be in a different state of recovery at the time of the test shock in the two conditions. In all our experiments the refractory periods were identical after both types of conditioning stimulus, and this suggests that if there were any refractory process, complete recovery must have occurred in less than 0·5 msec., since stimulus intervals as short as, or even shorter than, this have been observed after electrical conditioning stimuli. If, on the other hand, there were an intermediate structure not stimulated by the electrical conditioning shock, then complete recovery must have occurred before the absolute refractory period of the nerve was at an end.

The facilitation curve represents a change in excitability of the preparation, and it is difficult to conceive a mechanism which can account for these changes in excitability without involving the axon. Katz (1947) has shown in nerve that similar changes in excitability are accompanied by a change in the potential across the membrane, and it seems probable that the subthreshold mechanical stimulus sets up within the corpuscle a non-propagated potential. Such potentials have recently been recorded from frog's muscle spindles by Katz (1949). The exact point at which the action potential is set up in our facilitation experiments is unknown, and it may be some distance from the point of mechanical stimulation, in which case the facilitation curve represents a distorted picture of the time course of the subthreshold potential. It is evident that changes occur in the axon from the moment of the application of pressure to the corpuscle.

Whatever explanation is advanced to explain the initiation of an action potential or a series of action potentials by a mechanical stimulus, it must be able to account for the findings described above, and in particular for the following facts: the latency of the response to a threshold stimulus is 1·0 to 1·7 msec. and remains unaltered in the relatively refractory period; the latency has a minimum value around 0·5 msec.; the latency is little affected by the stimulus duration; the stimulus to the axon itself is graded and begins with the onset of the stimulus; and several action potentials can be produced by a stimulus of long duration.

It could be argued that the long latency is due to mechanical causes, i.e. damping or the natural period of the corpuscle. Damping might increase the rising time of a long stimulus to 1·5 msec., but could not do this to a stimulus of only 170 μsec. duration. If the stimulus started the corpuscle oscillating at its natural frequency, the threshold latency would be a quarter cycle. In this case the latency would approximate to a sine curve (of the form $y = a + b \sin^{-1} (1/x)$, where $y = $ latency and $x = $ stimulus strength) and large enough stimuli should reduce the latency to zero; neither of the conditions is fulfilled.

It is possible to postulate a mechanism intermediate between the mechanical stimulus and the excitation of the axon. For example, the corpuscle, rather than the axon inside it, might contain a means of producing a potential gradient. Such a potential would have to be initiated in less than 150 μsec.; have to remain high enough to maintain a charging current to the axon for 1·5 msec.; have a component maintainable by a constant deformation, and have a maximum value such that the minimal latency is about 0·5 msec. The same properties would have to be ascribed to any chemical intermediary, but there is strong evidence against the participation of acetylcholine or histamine in the initiation of action potentials at certain sensory endings (Brown & Gray 1948).

It is, in our opinion, quite unnecessary to postulate any intermediate mechanism at all, since the facts can be explained by the direct action of the stimulus on the axon. It has been shown by Schmitz & Schaefer (1933), Blair (1936), and Aird & Pfaffman (1947) that mechanical stimuli can excite nerves, and by Grundfest (1936) that pressure can increase their excitability. It is reasonable to assume that the mechanical stimulus results in a lowering of the membrane potential over a length of axon. If this length were longer than the critical length required for the propagation of an impulse, propagation would occur with only a short delay if the depolarization reached threshold; on the other hand, if this degree of depolarization were not reached during the short stimulus, further depolarization could not occur. However, this explanation cannot be correct, since it means that the longest latency would be of the same order as the stimulus duration, while in our experiments, we observed latencies of 1·5 msec. after 150 μsec. mechanical pulses. If a length shorter than the critical length were depolarized to a sufficient extent, the charge could spread to the surrounding axon, and raise a sufficient length to the threshold required for propagation; and the time for such a spread would account for the latencies observed. The required degree of the initial depolarization would be much reduced if the spread was into more excitable axon. An increase in stimulus strength would increase either the degree of depolarization or the area depolarized and so reduce the latency. It seems reasonable to suggest that the initial movement of the stimulator is a much more effective stimulus than its maintained pressure, and that this movement induces a sudden depolarization of the axon; a depolarization which declines according to the time constant of the membrane. The constant pressure can presumably, however, maintain some degree of depolarization, since a long stimulus of sufficient strength can produce a repetitive discharge. The comparatively long minimum latency always observed after maximum mechanical stimuli can be explained by assuming that there is a maximum length of axon susceptible to the mechanical stimulus.

We have shown that the absolute value of the latency after stimuli of just threshold level is the same at all points in the recovery curve. The refractory period of a local potential is shorter than the refractory period of a propagated potential (cf. Hodgkin 1938). If the difference between these two refractory periods is less than the latency then the local potential in the corpuscular axon would always be developing into axon which has recovered relative to the state of development of the local potential. Such an explanation could account for the remarkably

constant value of the latency after near threshold stimuli at a variety of levels of excitability of the corpuscle.

We feel that our findings can be explained adequately by assuming nothing other than the known properties of axons, and that the Pacinian corpuscle acts simply as a means of applying the mechanical stimulus to the axon.

We wish to thank Dr G. L. Brown, F.R.S., for his encouragement and advice throughout this work.

REFERENCES

Adrian, E. D. 1926 *J. Physiol.* **62**, 33.
Adrian, E. D. 1930 *Proc. Roy. Soc.* B, **106**, 596.
Adrian, E. D. & Umrath, K. 1929 *J. Physiol.* **68**, 139.
Aird, R. B. & Pfaffman, C. 1947 *Proc. Soc. Exp. Biol., N.Y.*, **66**, 130.
Andral, A. G. 1837 *Observations et proposition d'anatomie de chirurgie et de medicine.* Paris. Cited by *British and Foreign Medical Review*, 1845, **19**, 78.
Bernhard, C. G., Granit, R. & Skoglund, C. R. 1942 *J. Neurophysiol.* **5**, 55.
Blair, E. A. & Erlanger, J. 1938 *Amer. J. Physiol.* **121**, 431.
Blair, H. A. 1936 *Amer. J. Physiol.* **114**, 586.
Brown, G. L. & Gray, J. A. B. 1948 *J. Physiol.* **107**, 306.
Erlanger, J. & Gasser, H. S. 1937 *Electrical signs of nervous activity.* Philadelphia: University of Pennsylvania, Eldridge Reeves Johnson Foundation Lecture.
Gammon, G. D. & Bronk, D. W. 1935 *Amer. J. Physiol.* **114**, 77.
Gernandt, B. & Zotterman, Y. 1946 *Acta physiol. Scand.* **12**, 56.
Granit, R. 1947 *Sensory mechanism of the retina.* Oxford: University Press.
Gray, J. A. B. 1947 *J. Physiol.* **106**, 34*P*.
Gray, J. A. B. & Malcolm, J. L. 1949. *J. Physiol.* **108**, 43*P*.
Grundfest, H. 1936 *Cold Spr. Harb. Symp. Quant. Biol.* **4**, 179.
Henle, J. & Kölliker, A. 1844 *On the Pacinian corpuscles of the nerves of men and mammalia.* Zürich. Abstracted by *British and Foreign Medical Review*, 1845, **19**, 78.
Hodgkin, A. L. 1938 *Proc. Roy. Soc.* B, **126**, 87.
Katz, B. 1937 *Proc. Roy. Soc.* B, **124**, 244.
Katz, B. 1947 *J. Physiol.* **106**, 66.
Katz, 1949 *J. Physiol.* **109**, 9*P*.
Lehmann, J. G. 1741 *Dissertatio inauguralis medica de consensu partium corporis humani occasione spasmi singularis in manu eiusque digitis ex hernia observati; exposito simul nervorum brachialium et cruralium coalitu peculiari atque papillarum nervearum in digitis dispositione.* Vitembergae.
Matthews, B. H. C. 1933 *J. Physiol.* **78**, 1.
Pacini, F. 1835 *Proc. Florentine Med. Phys. Soc.* 22 November. Reprinted in *Sperimentale*, 1924, **78**, 275.
Pacini, F. 1840 *Micro organi scoperti nil corpo umano.* Pistoja.
Schmitz, W. & Schaefer, H. 1933 *Pflüg. Arch. ges. Physiol.* **232**, 7.
Skoglund, C. R. 1942 *Acta physiol. Scand.* **4**, Suppl. XII.
Todd. R. B. & Bowman, W. 1845 *The physiological anatomy and physiology of man.* London: J. W. Parker.

The form-transformation of the abdomen of the female pea-crab, *Pinnotheres pisum* Leach

By A. E. Needham

Department of Zoology and Comparative Anatomy, University Museum, Oxford

The growth in width (W) of the segments of the abdomen relative to carapace size (S), and the graded distribution of growth along the abdomen, are analyzed by the method of fitting, to the observed values of W, polynomial regressions of progressively higher power in S. The simplest (linear) relation reveals the main features and each closer approximation furnishes further detail.

The second object of the method, to select the lowest power of polynomial which adequately represents the data, gives the quadratic, though it is found that its adequacy varies in the different segments, which demand, for uniform adequacy, a non-affine set of polynomials. Adequacy is determined from the residual variance.

The set of quadratics for the seven segments of the abdomen are combined, by a modification of Medawar's transformation method, to give a single key relation which, within the scope of the data, defines abdomen width completely, spatially and temporally. This step involves the definition of the parameters of the quadratics as continuous functions of abdomen width at selected body size. It is suggested that the key relation to the transformation might, by analogy, be termed the 'form-cinematogram' for abdomen width. The equation: 'form = shape + size' is useful in the present context and is advocated for general recognition. The 'shape-cinematogram' may be derived from the form-cinematogram.

Alternative attempts to derive a satisfactory form-cinematogram from the data are outlined. The form change is surprisingly simplified by the excision of the initial width measurements from all subsequent width measurements.

The overall change in shape of the abdomen is visualized by the co-ordinate transformation method applied reciprocally between initial and final proportions.

Introduction

In a previous study of the growth of the abdomen of this crab (Williams & Needham 1938), inspired by preliminary observations by Huxley (1932, p. 94), the graph relating width of abdominal segment to carapace size was fitted by two successive straight lines (1938, text-figure 5) on arithmetic graph paper, and by three such lines on log/log paper (1938, text-figure 3). These arbitrary approximations were justified in providing a simple method of analysis of the temporal and spatial variations in the differential growth of the abdomen, but they contradicted the manifest smoothly continuous change in growth rate of the abdomen relative to the carapace. The 'growth constant', α (Huxley & Teissier 1936), was not constant, even over restricted periods of growth (an anomaly overlooked in Reeve & Huxley 1945), but was a continuously variable function of carapace size. The picture of growth then obtained was therefore disjointed, temporally. It was also disjointed spatially, into seven discrete width-measurements. The present analysis of the data aims at a more complete, algebraical, definition of the form of the abdomen, continuous both spatially and temporally. For such a continuous record, an alge-

braical motion-picture, of form change the term 'algebraical form-cinematogram'*
may be proposed, in the same sense as a cardiogram is the continuous record of
cardiac action, traced by a cardiograph. The algebraical method by which Medawar
(1944) defined, in a single key equation, the complete series of length proportions
of the human body as a continuous function, both of their initial proportions and of
time, has been adapted for this purpose.

METHOD

The data differ from those of Medawar (1944) since they are widths, measured at
intervals across the main axis, and not lengths measured along that axis. More-
over, body size—$\sqrt{}$(carapace length × carapace width)—has been substituted for
time (cf. Medawar 1944, p. 135), and the study is therefore, as previously (Williams
& Needham 1938), one of relative and not of absolute growth. The treatment of the
data also differs. In the first place absolute magnitudes have been used and not
relative proportions (which did not simplify the present problem (p. 131)); this
distinction may be given recognition by referring to Medawar's key equation
(Medawar 1945, p. 174) as a 'shape-cinematogram', in contradistinction to the
present form-cinematogram which expresses a (shape + size) transformation.†
Secondly, the two steps by which Medawar obtained his key relation are here
reversed. There, the series of length proportions at each age were related to the
series of initial proportions by a set of transformation relations the parameters
of which, after reduction to their efficient number, were then expressed as a con-
tinuous function of time. Here the width of each abdominal segment is related
to the 'time' dimension (carapace size) as the first step, and the parameters of the

---

* The term 'cinematogram' has been used by Fischer (1946, p. 22), without comment on
its origin, in the more orthodox sense of a record (? necessarily recorded by a cinematograph)
of the movement of cells in tissue culture.

† The distinction between 'shape' and 'form' here proposed and symbolized by the
equation form = shape + size, is probably novel. In current usage shape and form are treated
as virtually synonymous (e.g. Thompson 1942, pp. 22, 81, 286, 1026; chapter and running
headings; etc.) though the distinction proposed here is probably implicit in certain contexts
(Thompson 1942, pp. 206, 210). The present need for a distinction between shape and shape
+ size is therefore met by the resolution of an apparent tautology. The technical importance of
the distinction in the present instance is evident, and it clearly has a general and theoretical
value. Thus, for instance, *Pinnotheres* would not have received its popular name had its shape
only, and not its form also, resembled that of the culinary pea. So great is the ability of the
human eye to abstract shape from form, however, that the customary abstraction has rarely
been noticed or its recognition demanded. Students of shape usually reduce their forms to
a common overall size for purposes of illustration (Thompson 1942, chap. 17). As in the early
stages of any scientific investigation, students of growth and form have been interested in
analysis, that is, either in size or in shape, and no need for a terminological distinction between
shape and form has therefore been felt until the present type of situation is encountered, when
it is desired to synthesize shape change and size change into a complete definition of growth
(i.e. of form change). Interest in form change owes much to Huxley (1932), since his relation
of relative growth, $y = bx^\alpha$, deals with absolute dimensions, that is, with size and size change,
as well as with relative proportions (of x to y), that is, with shape and shape change. The
dynamic form of the above equation, viz. form change = (shape change) + (size change),
may clearly be written in the alternative form:

(differential growth) = (shape change + absolute growth).

set of relations so obtained are then expressed in terms of a standard set of width measurements.

No biological significance necessarily inheres the form of the equations used in the first step. Polynomial regression relations have been arbitrarily selected, largely because (but see p. 117 and figure 1) accurate statistical methods of fitting are available for this type of relation (Fisher 1946, p. 146). Otherwise it would have been reasonable to select the allometry relation $y = bx^{\alpha}$ of Huxley (Huxley 1932; Huxley & Teissier 1936) with α to be defined as a continuous function (p. 115) of the time dimension.

TABLE 1. COMPUTED VALUES OF THE PARAMETERS a, b, c, ... IN THE POLYNOMIAL REGRESSIONS $W = a + bS + cS^2$..., RELATING THE WIDTH (W) OF THE ABDOMINAL SEGMENTS OF THE FEMALE *PINNOTHERES PISUM* TO CARAPACE SIZE (S)

value of	segments of abdomen						telson
	1	2	3	4	5	6	
A. In the linear regression $W = a + bS$							
a	$-1\cdot90$	$-2\cdot35$	$-3\cdot91$	$-5\cdot04$	$-5\cdot24$	$-5\cdot00$	$-3\cdot71$
b	$+0\cdot84$	$+0\cdot99$	$+1\cdot39$	$+1\cdot67$	$+1\cdot70$	$+1\cdot57$	$+1\cdot12$
B. In the quadratic regression $W = a + bS + cS^2$							
a	$-4\cdot20$	$-5\cdot27$	$-7\cdot71$	$-11\cdot22$	$-12\cdot20$	$-11\cdot57$	$-8\cdot51$
b	$+1\cdot39$	$+1\cdot69$	$+2\cdot30$	$+3\cdot15$	$+3\cdot37$	$+3\cdot14$	$+2\cdot27$
c	$-0\cdot032$	$-0\cdot040$	$-0\cdot052$	$-0\cdot084$	$-0\cdot095$	$-0\cdot090$	$-0\cdot066$
C. In the cubic regression $W = a + bS + cS^2 + dS^3$							
a	$-4\cdot21$	$-8\cdot03$	$-12\cdot00$	$-16\cdot40$	$-17\cdot01$	$-13\cdot14$	$-6\cdot32$
b	$+1\cdot40$	$+2\cdot71$	$+3\cdot89$	$+5\cdot07$	$+5\cdot15$	$+3\cdot72$	$+1\cdot46$
c	$-0\cdot032$	$-0\cdot16$	$-0\cdot24$	$-0\cdot31$	$-0\cdot30$	$-0\cdot16$	$+0\cdot030$
d	$+0\cdot00001$	$+0\cdot0045$	$+0\cdot0070$	$+0\cdot0084$	$+0\cdot0079$	$+0\cdot0026$	$-0\cdot0036$
D. In the quartic regression $W = a + bS + cS^2 + dS^3 + eS^4$							
a	$-4\cdot52$	$+2\cdot55$	$-11\cdot48$	$-3\cdot72$	$-2\cdot57$	$-4\cdot61$	$-5\cdot83$
b	$+1\cdot55$	$-2\cdot61$	$+3\cdot63$	$-1\cdot31$	$-2\cdot12$	$-0\cdot57$	$+1\cdot22$
c	$-0\cdot060$	$+0\cdot80$	$-0\cdot19$	$+0\cdot85$	$+1\cdot01$	$+0\cdot62$	$+0\cdot074$
d	$+0\cdot0022$	$-0\cdot070$	$+0\cdot0031$	$-0\cdot081$	$-0\cdot094$	$-0\cdot058$	$-0\cdot0070$
e	$-0\cdot0000038$	$+0\cdot0021$	$+0\cdot00011$	$+0\cdot0025$	$+0\cdot0029$	$+0\cdot0017$	$+0\cdot000096$

Irrespective of any physiological significance in the form of the selected relations, however, the variation of the parameters of the set of polynomial relations for the seven segments (table 1) may be used as an index of the variation in growth pattern along the abdomen (figure 2). The adequacy of fitting of the polynomials (figure 1) is in fact very good, and it would seem reasonable to go further and to use the first differential as an adequate measure of growth rate of the segment relative to carapace size, the second derivative as a measure of rate of change in relative growth rate, and so on. The analysis of the relations on these lines forms the first section of the study (pp. 120 to 127).

The exercise of applying Fisher's method may itself be regarded as an important part of the present work, and polynomials of progressively higher power have been fitted in order to ascertain, from a scrutiny of the residual variance (table 3),
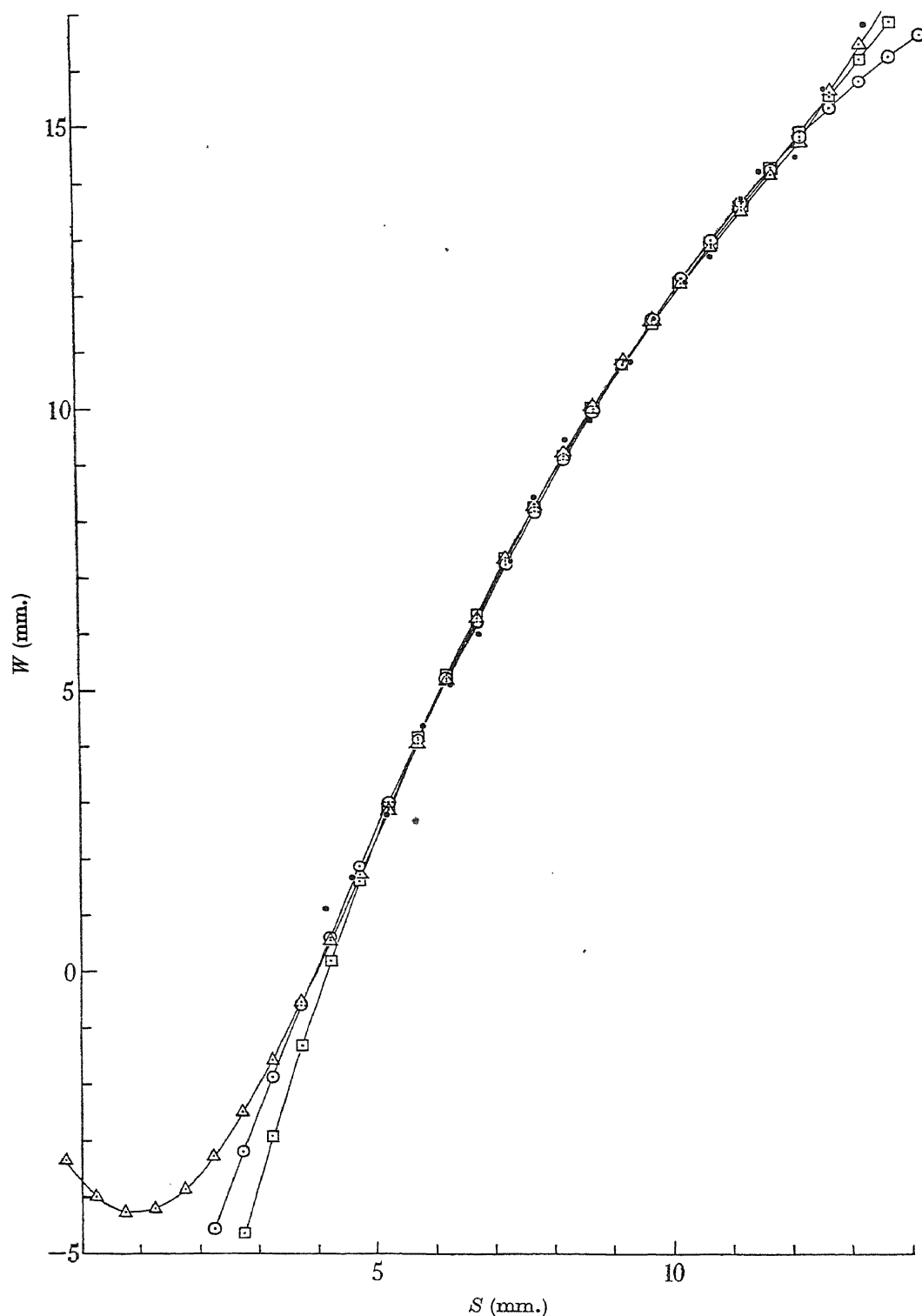
FIGURE 1. Graphical representation of the relation between the width (W) of the fourth abdominal segment of the female *Pinnotheres pisum* and carapace size (S). The solid circles give the observed data (grouped as in Williams & Needham 1938, table 1). For comparison the open circles give the values of W, at intervals of 0·5 mm. of S, computed from the best quadratic polynomial $W = a + bS + cS^2$ fitted to the observed data. The squares and triangles give the corresponding values computed from the best cubic and quartic polynomials respectively. The curves obtained from the polynomials are extended beyond the range for which observed data were available, in order to indicate their turning points and other features. Arithmetic graph paper.

the lowest power which adequately represents the observed data. This progressive method of fitting also permits a logical analysis of the growth picture, the simplest relation showing merely the broadest features, and each succeeding power adding further detail. For this reason the linear regression has been included as the first step, although it is manifest, from the curvature of the graphs (Williams & Needham 1938, text-figure 5), that a single straight line is grossly inadequate.

TABLE 2. POLYNOMIAL VALUES OF W, OBTAINED FROM THE QUARTIC REGRESSION RELATIONS IN S

(Compare with observed values: Williams & Needham 1938, table 1.)

mean carapace size S (mm.)	width (W) of abdominal segments (mm.)						
	1	2	3	4	5	6	telson
13·25	8·74	10·56	14·06	16·40	16·40	14·61	9·94
12·75	8·46	9·98	13·42	15·54	15·56	14·02	9·72
12·25	8·16	9·47	12·80	14·79	14·83	13·46	9·46
11·75	7·84	9·03	12·20	14·10	14·16	12·92	9·14
11·25	7·50	8·62	11·60	13·46	13·53	12·37	8·78
10·75	7·15	8·23	11·01	12·82	12·91	11·80	8·37
10·25	6·79	7·84	10·40	12·17	12·28	11·19	7·93
9·75	6·40	7·43	9·79	11·50	11·60	10·55	7·44
9·25	6·00	7·00	9·16	10·78	10·88	9·85	6·91
8·75	5·59	6·54	8·51	10·00	10·09	9·10	6·35
8·25	5·16	6·04	7·83	9·16	9·22	8·30	5·75
7·75	4·72	5·50	7·11	8·26	8·29	7·43	5·13
7·25	4·26	4·93	6·36	7·29	7·29	6·51	4·48
6·75	3·78	4·32	5·57	6·26	6·22	5·54	3·80
6·25	3·29	3·70	4·72	5·17	5·09	4·52	3·10
5·75	2·78	3·07	3·83	4·04	3·92	3·47	2·39
5·25	2·25	2·44	2·87	2·89	2·75	2·41	1·66
4·75	1·71	1·83	1·86	1·73	1·55	1·34	0·92
4·25	1·15	1·27	0·79	0·58	0·39	0·28	0·17

Curves empirically fitted to the observed data show clear evidence of two inflexions, one near the beginning of the size range (4 to 13·5 mm. carapace size, table 2) and the other towards the end, so that a quartic expression in carapace size (S) would seem necessary to define adequately abdomen width (W). Technically the labour of computation was somewhat reduced by reversing the logical progression and fitting a set of quartic relations as the first step, and the lower powers, cubic, quadratic and linear, in order by the progressive omission of the highest term. Analysis of the variance (table 3) indicated that no power higher than the quartic could usefully be applied (p. 125). The polynomial values corresponding to the fitting of each power of polynomial were computed as suggested (Fisher 1946, p. 172) from the polynomial values at the working mean, together with its progressive differences. Those for the quartics, only, are reproduced in table 2. In comparing them with the observed mean group values of W (Williams & Needham 1938, table 1), the inequalities in the intervals between successive observed group values of S must be borne in mind. The W/S curves plotted from the computed polynomial values, for segment 4 only, are shown in figure 1, superimposed upon

the observed data. The three curves, quartic, cubic and quadratic, are seen to differ little except towards the limits of the size range of the data, where records were few, and all fit the observed data very closely, so that, in general, no great loss of precision would result from the application of any one of the three.

The equations to the fitted W/S relations were obtained from the computed polynomial values by the usual algebraical method, and are given in table 1. Real solutions of the equations $W = 0$, maxima and minima of W (i.e. solutions of the equations $dW/dS = 0$), turning points (solutions of $d^2W/dS^2 = 0$), and other significant characteristics of the relations, are shown in table 4.

RESULTS

I. *Analysis of the polynomial regression relations*

(1) *The linear regression:* $W = a + bS$

This shows the main features of relative growth in the abdomen (table 1 A). The coefficient b, the slope of the regression line, gives a measure of the mean rate of relative growth in segment width over the range of body size covered by the data. This measure is, of course, not only very approximate, but also purely empirical (p. 117), and is not to be accorded the significance of measures of relative growth rate which are backed by theoretical considerations, as, for instance, the growth index of Huxley (p. 115). The value of b is greater than unity except in segments 1 and 2. In all segments the ratio of W to S increases with S, since a is negative in sign. The relative growth rate is maximal in segment 5 (figure 2) and declines progressively both distalwards and proximalwards. As pointed out previously (Williams & Needham 1938, p. 549) the growth centre appears to be farther forward on linear than on logarithmic plotting (cf. also Huxley 1932, p. 95). The co-ordinate transformation (figure 5) naturally shows the linear picture of the growth gradient.

The parameter a gives, by extrapolation, the value of W in each segment when $S = 0$. Its negative sign indicates that the curve of growth subsequent to $S = 4$ mm. does not continue an earlier curve but follows on, via a relatively sharp discontinuity (cf. p. 132), from the earlier phase of low relative growth rate. The arithmetical magnitude of a is mainly a reflexion of that of b (of opposite sign). The two parameters are approximately related to each other by a straight line which does not pass through the origin. More precisely the relation is elliptical (p. 124). The ratio $(-a/b)$ gives the value of S when $W = 0$. This value increases 'monotonically' (Medawar 1945, p. 177) from the proximal to distal end of the abdomen (table 4), and may conceivably indicate that the phase of growth studied has a progressively later time of onset in the proximo-distal direction. This gradient contrasts (cf. p. 133) with the 'dyatonic' (Medawar 1945, p. 177) gradient in average growth rate (p. 120).

Although the linear regression so inadequately fits the observed data (p. 119), it is noteworthy (table 3) that it accounts for 98 % of the total variance between classes, after deducting that due to the position of the working mean. The magni-

TABLE 3. ANALYSIS OF THE VARIANCE, BETWEEN CLASSES, OF THE WIDTH (W) OF THE ABDOMINAL SEGMENTS OF THE FEMALE *PINNOTHERES PISUM*, FITTED BY POLYNOMIALS IN BODY SIZE (S)

order of polynomial in S ...	linear	quadratic	cubic	quartic
Segment 1. Total variance between classes = +560·33				
variance due to position of working mean	+ 97·57	+ 236·54	+ 236·46	+ 234·78
variance due to term in: S	+ 451·65	+ 427·59	+ 427·54	+ 428·04
S^2	—	− 110·87	− 110·76	− 108·45
S^3	—	—	− 0·04	+ 0·90
S^4	—	—	—	− 2·14
residual variance	+ 11·09	+ 7·07	+ 7·13	+ 7·16
Segment 2. Total variance between classes = +773·24				
variance due to position of working mean	+ 132·93	+ 335·93	+ 326·92	+ 391·70
variance due to term in: S	+ 628·88	+ 592·94	+ 570·27	+ 548·18
S^2	—	− 160·70	− 111·90	− 201·84
S^3	—	—	− 16·60	− 51·28
S^4	—	—	—	+ 82·55
residual variance	+ 11·43	+ 5·07	+ 4·57	+ 3·93
Segment 3. Total variance between classes = +1522·42				
variance due to position of working mean	+ 245·99	+ 593·71	+ 575·27	+ 579·36
variance due to term in: S	+1260·19	+1193·54	+1143·28	+1140·86
S^2	—	− 270·30	− 172·16	− 177·34
S^3	—	—	− 28·19	− 30·48
S^4	—	—	—	+ 5·31
residual variance	+ 16·24	+ 5·46	+ 4·21	+ 4·70
Segment 4. Total variance between classes = +2206·04				
variance due to position of working mean	+ 342·00	+ 999·36	+ 973·59	+1092·09
variance due to term in: S	+1823·78	+1692·65	+1619·57	+1574·01
S^2	—	− 497·67	− 364·10	− 521·53
S^3	—	—	− 32·88	− 76·73
S^4	—	—	—	+ 129·75
residual variance	+ 40·27	+ 11·71	+ 9·92	+ 9·02
Segment 5. Total variance between classes = +2285·42				
variance due to position of working mean	+ 349·28	+1091·92	+1068·75	+1203·67
variance due to term in: S	+1888·10	+1737·90	+1668·63	+1615·60
S^2	—	− 556·29	− 432·75	− 611·59
S^3	—	—	− 28·60	− 76·89
S^4	—	—	—	+ 143·82
residual variance	+ 48·04	+ 11·89	+ 10·39	+ 10·82
Segment 6. Total variance between classes = +1944·14				
variance due to position of working mean	+ 291·83	+ 926·36	+ 919·18	+ 991·70
variance due to term in: S	+1611·80	+1480·69	+1459·75	+1430·87
S^2	—	− 471·27	− 435·02	− 529·56
S^3	—	—	− 7·99	− 6·59
S^4	—	—	—	+ 2·91
residual variance	+ 40·51	+ 8·36	+ 8·22	+ 7·85
Telson. Total variance between classes = +996·47				
variance due to position of working mean	+ 146·52	+ 472·21	+ 479·20	+ 481·83
variance due to term in: S	+ 827·67	+ 758·84	+ 779·74	+ 778·58
S^2	—	− 239·63	− 274·70	− 278·43
S^3	—	—	− 7·49	− 6·59
S^4	—	—	—	+ 2·91
residual variance	+ 22·26	+ 5·03	+ 19·71	+ 18·17

tudes of the total variance between classes, of the variance due to linear regression, and of the residual variance, all parallel that of a, b, having their maxima in segment 5 and a progressive decline proximalwards and distalwards.

(2) *The quadratic regression:* $W = a + bS + cS^2$

The decrease in residual variance (table 3), on fitting a quadratic (table 1B) in place of the linear relation, is considerable in most segments. In segment 5 it is decreased by more than three-quarters. Further, the residual variance (r.v.) becomes more nearly equal in all segments, indicating that it is now due chiefly to errors of measurement and to similar factors which do not differ greatly from segment to segment. The decrease in r.v. is roughly proportional to the relative growth rate, being maximal in segment 5 and declining dyatonically, i.e. in both directions (p. 120). Segment 1 appears anomalous; here the decrease in r.v. is small, so that r.v. is now greater than in segments 2, 3 and telson, and is almost as great as that of segment 5. In fact, greater difficulty was experienced in measuring this than any other segment, and the high r.v. may be due to this cause.

The effect of the term in S^2 on the total variance between classes is, of course, far greater than the decrease in r.v., and, indeed, the term contributes a large negative fraction to the total variance (table 3), which is, however, more than counterbalanced by an increased positive contribution due to the position of the working mean. The term in S makes a smaller positive contribution than on linear fitting.

In the quadratic relations (table 1B) the parameter a, as before, is the value of W when $S = 0$, again negative in sign and unreal, and reflecting the high positive value of b, which is now the relative growth rate at the outset (i.e. the value of dW/dS when $S = 0$). The parameter c measures the rate of change in relative growth rate and is invariant in S ($d^2W/dS^2 = 2c$). All three parameters are arithmetically (i.e. ignoring sign) greatest in segment 5 (cf. p. 120), with the familiar dyatonic gradient towards the ends of the abdomen; the magnitude of the change in growth rate is thus proportional to the magnitude of the initial growth rate. The sign of c is negative, expressing the manifest decline in relative growth rate with increase in S. As on linear fitting b is approximately one-third the magnitude of a; c is approximately one-fortieth the magnitude of b.

The quadratic relations give two real solutions to the equations

$$W = a + bS + cS^2 = 0.$$

One value lies around $S = 4$ mm., and is lowest in segment 1, with a monotonic increase distalwards to the telson (table 4); the other value lies between 30 and 41 mm. and is much more variable. Its magnitude shows a monotonic gradient in the reverse direction, the highest value being in segment 1. These monotonic gradients resemble that in the ratio a/b of the linear relations (p. 120). The quadratic relations are parabolae with their axes parallel to the W-axis; the value of S on the axis of the parabola (table 4) increases monotonically from the telson to segment 1, and it is therefore clear that the higher of the two values of S which put $W = 0$ shows the same gradient and that the parabolae for anterior segments are in general

more obtuse than those of posterior segments; in general, but not in strict order, since the length of the latus rectum does not follow the monotonic gradient (table 4), the value of W at the focus being maximal in segment 4, and its gradient along the

TABLE 4. CHARACTERISTICS OF THE POLYNOMIAL RELATIONS

(All measurements in mm.)

A. Quadratic

segment number	real values of S making $W=0$	co-ordinates of vertex of parabola		value of W at focus of parabola	length of latus rectum	values of W when $S=1{\cdot}0$
		S	W			
1	3·26, 41·02	22·14	11·24	3·30	31·75	−2·84
2	3·39, 39·07	21·23	12·67	6·35	25·28	−3·62
3	3·65, 40·81	22·23	17·90	13·07	29·30	−5·45
4	3·99, 33·37	18·68	18·23	15·28	11·84	−8·15
5	4·10, 31·36	17·73	17·64	15·01	10·54	−8·92
6	4·19, 30·87	17·53	15·96	13·17	9·60	−8·51
telson	4·28, 30·40	17·34	11·19	7·36	15·33	−6·30

B. Cubic

segment number	real values of S making $W=0$	values of S making $dW/dS=0$	values of S making $d_2W/dS^2=0$
1	$\begin{cases} 3{\cdot}25 \\ 41{\cdot}70 \\ 2555{\cdot}70 \end{cases}$	$\begin{cases} 22{\cdot}30 \ (W \text{ maximum}) \\ 1711{\cdot}50 \ (W \text{ minimum}) \end{cases}$	866·88 (dW/dS a minimum) —
2	3·60	none	11·14 (dW/dS a minimum)
3	3·92	none	11·32 (dW/dS a minimum)
4	4·18	none	12·19 (dW/dS a minimum)
5	4·22	none	12·88 (dW/dS a minimum)
6	4·24	none	20·47 (dW/dS a minimum)
telson	$\begin{cases} 4{\cdot}32 \\ 20{\cdot}10 \end{cases}$	14·75 (W maximum) —	1·38 (dW/dS a minimum) —

C. Quartic

segment number	real values of S making $W=0$	values of S making $dW/dS=0$	values of S making $d_2W/dS^2=0$
1	3·27	20·0 (W maximum)	none
2	2·25	2·2 (W minimum)	5·92 (dW/dS max.), 10·78 (dW/dS min.)
3	4·10	negative	−25·28 (dW/dS max.), 11·20 (dW/dS min.)
4	3·95	0·9 (W minimum)	5·11 (dW/dS max.), 11·02 (dW/dS min.)
5	3·90	1·27 (W minimum)	5·26 (dW/dS max.), 10·91 (dW/dS min.)
6	4·10	0·25 (W minimum)	5·13 (dW/dS max.), 11·87 (dW/dS min.)
telson	3·90	$\begin{cases} \text{very high } (W \text{ minimum}) \\ 15{\cdot}02 \ (W \text{ maximum}) \end{cases}$	3·91 (dW/dS max.), 35·79 (dW/dS min.)

abdomen dyatonic. The values of S putting $W=0$ are thus not alone sufficient to characterize the conics, which neither are coaxial nor have a common directrix.

These parabolae imply a maximum in W at their apices, and a growth rate declining to zero there $(dW/dS \to 0)$. The values of W and S corresponding to these maxima are also given in table 4. The half of the curve beyond the maximum, and the second, higher value of S putting $W=0$, must be regarded as biologically

meaningless, and even an eventual decline in relative growth rate to zero goes beyond available data. It is probably never observed in nature. A quadratic implies no inflexions ($d^2W/dS^2 \neq 0$), and therefore represents only the main features of the curves (p. 119). The inflexion near the upper end of the observed curves thus reverses the tendency towards zero relative growth rate inherent in the quadratic.

(3) *The cubic regression:* $W = a + bS + cS^2 + dS^3$

In segments 1 and telson the r.v. is greater (table 3) on fitting a cubic in S than for the quadratic, which must therefore be regarded as the best polynomial for these two segments. For all others there is a further decrease in r.v., though this is much smaller than that effected by quadratic over linear fitting. As before the decrease in r.v. is greatest in the segment of maximal growth rate (segment 5, figure 2) and diminishes dyatonically towards the ends of the abdomen, so that the increases in segments 1 and telson may conveniently be regarded as negative decreases. The increase in r.v. is very great in the telson, and certainly for this segment any power of polynomial higher than the quadratic must be rejected. Thus it would appear that a higher power is required for the more rapidly growing segments in the centre of the abdomen than for the terminal segments, and that the most adequate set of relations would *not be strictly affine* throughout the abdomen.

The contribution of the term in S^3 to the variance is negative, like that of S^2.

In the cubic relations (table 1C) the parameters a, b, c retain the same character and significance as in the quadratics, though their arithmetic magnitude is substantially greater. The gradients, along the series of segments, in the magnitude of each parameter become relatively steeper in the order a, b, c, d so that in the telson the sign of c, d is actually reversed, compared with that of other segments. The positive sign of d in typical segments expresses that check to the decline in relative growth rate which is indicated by the inflexion near the upper end of the observed curves (p. 119). On cubic fitting this check is invariant in $S(d^3W/dS^3 = a$ constant $= 6d)$. The arithmetic magnitude of d continues the decline $a \rangle b \rangle c$, and the ratios a/b, b/c, c/d increase in that same order, that is to say, the geometric rate of decrease in magnitude of parameters, along the series a, b, c, d, in any segment, increases in that order. Thus $a/b = 3, b/c = 17, c/d = 36$. The arithmetic rate of decrease, of course, shows the reverse sequence $(a - b \rangle b - c \rangle c - d)$. The parameters c, d have their maximal value in segment 4, anterior to that of a, b, so that, if it be supposed that the main phase of growth of the abdomen is associated with a 'centre' of growth in segment 5 (figure 2), then it may possibly be inferred that the subsequent check on growth rate and the countercheck are centred progressively farther forward in the structure. This proximal movement of the maximum for successive parameters results in elliptical curves when the parameters are plotted against each other (figure 3); this also appears in the quadratic and even in the linear relations.

Except in segments 1 and telson the cubic relations give only one real solution to the equation $W = a + bS + cS^2 + dS^3 = 0$, i.e. only one real value of S for which $W = 0$, a feature which is biologically correct. Again, except in segments 1 and telson, there is no real solution to the equation $dW/dS = 0$, i.e. no maximum or

minimum value of W, again biologically reasonable. The upper inflexion of the observed curves is given by the single real solution of $d^2W/dS^2 = 0$ (table 4). In segments 2 to 5 the inflexion lies within the size range of the data; for other segments it lies outside the range and may have no biological significance. The complete anomaly of segments 1 and telson is no doubt correlated with the superfluity of cubic fitting for these segments (p. 124).
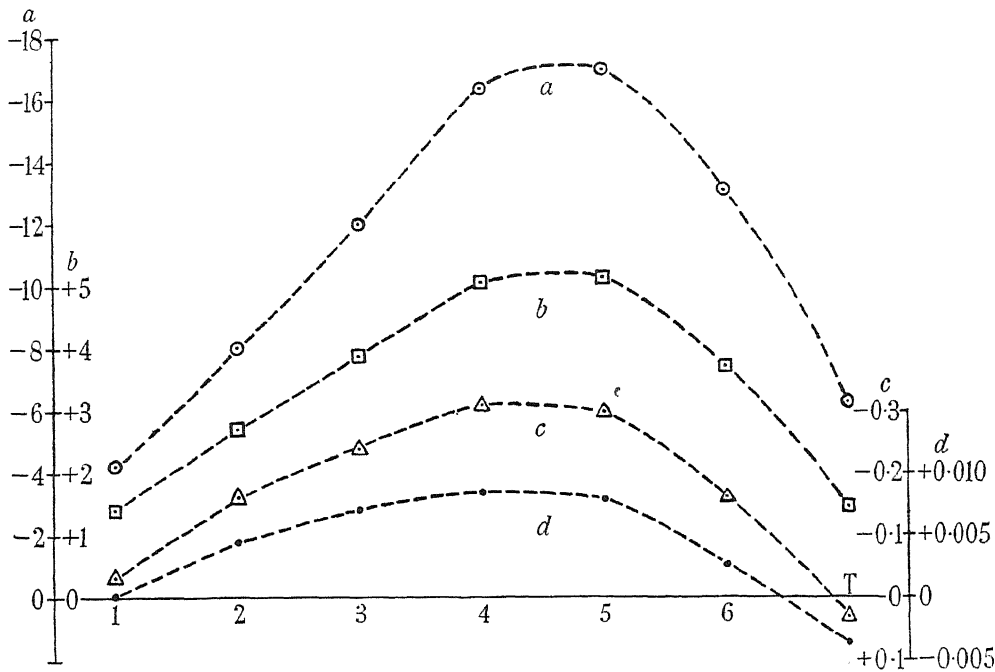
FIGURE 2. The four parameters a, b, c, d of the cubic polynomial relations (table 1C)

$$W = a + bS + cS^2 + dS^3,$$

relating width of abdominal segment (W) to carapace size (S), plotted against segment number, to show the gradients, along the abdomen, in the magnitude of each parameter. The ordinates give the magnitudes of the parameters, each to a separate scale, and the abscissa gives the number of the segment. For a, c, negative values are plotted upwards.

(4) *The quartic regression:* $W = a + bS + cS^2 + dS^3 + eS^4$

The fitting of quartics in S still further reduces the r.v. in segments 2, 4 and 6, i.e. curiously sporadically among the series of segments. In segment 2 alone is the reduction large, compared with cubic fitting, and the final r.v. here is probably sufficiently small to justify the assertion that a quartic adequately represents the data. Unfortunately, the original measurements on individuals are no longer available for an estimate of the variance between classes from which a certain test of adequacy could be made (Fisher, 1946, p. 172). In segments other than 2 (and, of course, 1 and telson (p. 124)) still higher powers would probably be required for precise fitting (cf. figure 1); however, the data scarcely justify any attempt to fit accurately the data near the extremes of the size range, and the quartic probably represents the utmost limit of utility within that range.

The r.v. of segment 1 remains high relative to other characteristics of the segment (cf. p. 122) and that of the telson is outstandingly high through 'over-fitting' (p.124). Except in segment 1 the contribution of the term in S^4 to the variance is positive, in contrast to those of S^2 and S^3 (table 3).

The table of parameters (table 1D) reveals a number of irregularities contingent on the more detailed fitting of the data, though the various gradients in magnitude and the alternation in sign along the series a, b, c, d, e are preserved in essence.

The quartic relation admits the inflexion near the lower end of the empirical curves (p. 119), and this leads to a minimum in W, by extrapolation back towards $S = 0$ (figure 1). Prior to this minimum the slope of the curve (i.e. the parameter b) necessarily takes a negative value. The values of a, however, are negative (unreal) still, as on simpler fitting, and the regular alternation of sign in the terms of the polynomial is disturbed. The parameter c becomes positive in sign and now measures the acceleration in growth rate leading to the high rate immediately above $S = 4$ mm. Its magnitude is increased correspondingly. Similarly, d, e effectively replace c, d respectively of the cubic, and b, c, d, e show the same regular alternation of sign as the parameters of the cubic.

In each segment the magnitude of the parameters (ignoring sign) decreases in the order a, b, c, d, e and the geometric rate of decrease increases in the same order (cf. p. 124). Again, the magnitude of each parameter, along the series of segments, decreases from its maximum in segment 5 towards each end of the abdomen, the gradient steepening towards the terminal segments so that in the anterior segments (segment 2 is anomalous) the sign of the parameter is opposite to that in segment 5. On cubic fitting the gradient appeared to be steeper towards the posterior end (see p.124 —reversal of sign in telson). In contrast to the parameters of the cubics all are maximal in segment 5, and the speculation (p. 124) that there may be separate centres of growth retardation, etc., must be treated with caution. Nevertheless, the polynomial values for segment 4 overtake those of segment 5, at maximal body size (table 2). The graphs of one quartic parameter against another give a rather irregular relation in place of the fairly regular ellipse of the lower polynomials.

This irregularity in the magnitude of the parameter, along the series of segments, increases progressively from e, which has relatively smooth gradients, to a. Detailed differences between segments begin to obscure the more fundamental features. It is possible that the quartics are beginning to fit sampling and other 'errors' in the material.

Quartic fitting makes 2 the widest segment prior to $S = 4$ mm. which is in harmony with available data on that range of body size (Lebour 1928; Sandoz & Hopkins, 1947). Segment 3 is also anomalous, and it is perhaps significant that this is the widest segment in the male (figure 5B); here, once more (cf. Needham 1937, p. 310), a male feature may affect the female, in vestigial form.

Quartics, like the cubics, give only one real value of S for $W = 0$ (table 4), and it has the same form of gradient in magnitude along the series of segments. They give a minimum value of W between $S = 0$ and $S = 4$ in segments 2, 4, 5 and 6, and a maximum in 1 and telson, at 20 and 15 mm. respectively. In the telson there is

a minimum at a value of S beyond the range studied. The positions of the two inflexions of the curves are very constant in S in the more typical segments, 2 to 6, i.e. more constant than any of the characteristics of the quadratic (p. 122) or of the cubic (table 4). In this respect, therefore, quartic fitting would seem to be the most adequate. Segment 1 has no inflexions, and those of 3 and telson are anomalous; in this respect, again, quartic fitting is unsatisfactory for practical purposes.

Thus it may be concluded that over the size range of the present data a quartic in S is probably more than adequate to define W in any segment and to extract all useful information from the data. In the terminal segments no power higher than the quadratic is legitimate. The fact that different powers are required for the different segments diminishes the probability that the relations can have more than empirical value, though their empirical value, as analytical tools, is seen to be considerable. The choice of power of polynomial is, to some extent, a matter of convenience, depending on the nature of the information required. The quadratic adequately describes the main features of relative growth in the abdomen over this size range, and specifically eliminates the mathematical complexity due to phases of growth which are largely outside this range. It has therefore been selected for the derivation of the form-cinematogram.

II. *The form-cinematogram*

The demand (p. 115) for an 'algebraical form-cinematogram', that is to say, a key relation (Medawar 1945, p. 174) which shall define the width at any point along the axis of the abdomen, throughout the size range studied, in terms of body size and of the value of the width at that point at some appropriate 'standard' body size, can be satisfied by defining each of the parameters $a, b, c, \ldots$ of the accepted set of polynomials in S (table 1) as a continuous function of abdomen width at the standard body size. Ideally, only one such standard abdomen size should be used, all the parameters of the polynomials being therefore expressed in terms of this one 'efficient' parameter, and the practical value of the cinematogram deteriorates in proportion to the number of efficient parameters (Medawar 1945, p. 175). However, the complexity of the form change in a structure will determine the minimum number of efficient parameters necessary for its definition. The present example is certainly not simple, since the position of the segment of maximum width moves back from an anterior segment (first or second according to the power of polynomial fitted (table 2)) at the outset ($S = 4$ mm.) to segment 5, where it remains until near the end of the period studied, and then moves forward again to segment 4. Moreover, prior to $S = 4 \cdot 0$ (Atkins 1926; Sandoz & Hopkins 1947) proportions have already changed radically at least once, and the maturing female abdomen therefore builds on a foundation of infantile proportions (or disproportions! cf. p. 132). The effect of 'excising' the infantile portion of the abdomen is striking. The sharpness of the increase in relative growth rate around $S = 4 \cdot 0$ virtually amounts to a discontinuity in growth, which can scarcely be defined algebraically. Clearly a complete cinematogram, of some complexity, could theoretically be obtained, but the value of complex relations is limited, at present, and the immediate aim is the

definition of a relatively simple phase of growth by a relation which is not too unwieldy for further use.

With this aim the quartic and cubic relations must be rejected in the first instance. The parameters of the former (table 1), while generally related (p. 126) to such characteristics of the segment as its width, W_0 or $W_{13 \cdot 25}$, at $S = 0$ or $S = 13 \cdot 25$,
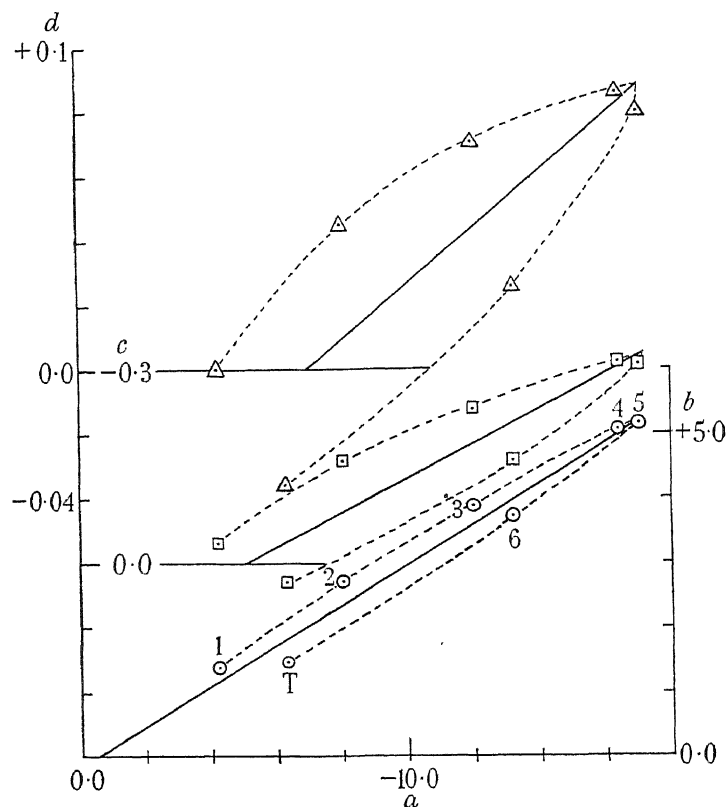
FIGURE 3. Graphical representation (arithmetic paper) of the interrelationships of the four parameters a, b, c, d of the set of cubic polynomials (table 1C) relating width (W) of abdominal segment with carapace size (S). The parameters b, c, d are plotted as ordinates against a as abscissa. The letters b, c, d are placed on their appropriate scale. On the scale for c negative values are plotted upwards and on that for a negative values are plotted to the right. The number of the segment is indicated, for all three graphs, beside the points on the b/a graph. The relations are seen to approximate to portions of ellipses.

respectively, vary too much for any very simple definition. The parameters of the cubic relations are more smoothly graded in magnitude along the abdomen, but even here definition would be very complex. As already noted the parameter a is itself identical with W_0. The graphs of the other parameters b, c, d, against a (i.e. against W_0), for the cubic relations approximate to ellipses (figure 3) of which the proximal and distal groups of segments each contribute one-half, about the major axis, but since neither apex of the ellipse lies at the origin, a simple single-valued definition of these parameters in terms of W_0 is not possible. The elliptical form is inevitable since the position of the maximum value of b, c, d moves progressively farther forward in the abdomen (p. 124).

As an approximation two separate quadratics in a may be fitted respectively to proximal and distal groups of the values of b, c, d (table 5). This procedure is perhaps justified by the manifest differences between the two groups of segments, as well as by the simplification effected, but quadratic relations each fitted to only three or four points are necessarily limited in value. The procedure has the further merit, however, of defining all four parameters in terms of a single standard series of measurements, namely, W_0. In table 5 the corresponding dual definition of the parameters b, c of the quadratic regressions is also given. The key equation is

TABLE 5. THE PARAMETERS b, c OF THE QUADRATIC REGRESSION RELATION (SEE TABLE 1), AND THE PARAMETERS b, c, d OF THE CUBIC REGRESSION, EXPRESSED AS QUADRATIC FUNCTIONS OF THE PARAMETER a, USING SEPARATE RELATIONS FOR PROXIMAL AND DISTAL GROUPS OF SEGMENTS

A. Quadratic

proximal

distal

$b = +0\cdot17 - 0\cdot31a - 0\cdot0037a^2$

$b = +2\cdot07 + 0\cdot17a + 0\cdot022a^2$

$c = -0\cdot0062 + 0\cdot0051a - 0\cdot00016a^2$

$c = -0\cdot019 + 0\cdot0037a - 0\cdot00021a^2$

B. Cubic

$b = -0\cdot19 - 0\cdot40a - 0\cdot0047a^2$

$b = -0\cdot34 - 0\cdot26a + 0\cdot0035a^2$

$c = +0\cdot14 + 0\cdot045a + 0\cdot0011a^2$

$c = +0\cdot12 + 0\cdot0085a - 0\cdot00097a^2$

$d = -0\cdot0065 - 0\cdot0018a - 0\cdot000053a^2$

$d = -0\cdot0056 - 0\cdot000043a + 0\cdot000044a^2$

TABLE 6. VALUES OF W WHEN $S = 10$ MM., DERIVED DIRECTLY FROM THE POLYNOMIAL RELATIONS FOR INDIVIDUAL SEGMENTS, COMPARED WITH THE VALUES DERIVED FROM THE SINGLE CINEMATOGRAM OF PAGE 131 AND WITH THE VALUES DERIVED FROM THE SEPARATE PROXIMAL AND DISTAL CINEMATOGRAMS OBTAINED FROM TABLE 5

A. Quadratic regression

segment number	values of W_{10} obtained from the regression relations of individual segments (mm.)	values of W_{10} obtained from the single cinematogram of p. 131 (mm.)	values of W_{10} obtained from the dual cinematogram of table 5 (mm.)
1	6·596	6·836	6·644
2	7·653	7·903	7·778
3	10·147	9·721	9·914
4	11·878	11·790	11·880
5	11·976	12·084	11·914
6	10·876	10·873	10·883
telson	7·657	7·573	7·656

B. Cubic regression

segment number	values of W_{10} obtained direct from regression relations to individual segments (mm.)	values of W_{10} obtained from the dual cinematograms of table 5 (mm.)
1	6·597	6·633
2	7·622	7·741
3	10·099	9·441
4	11·839	11·672
5	11·921	11·903
6	10·862	10·845
telson	7·686	7·711

A. E. Needham

obtained for proximal and distal halves of the abdomen separately, by substituting the appropriate expressions in W_0 for a, b, c, d in the general equation

$$W_S = a + bS + cS^2 + dS^3.$$

The adequacy of these key relations may be tested by comparing the segmental values of W for $S = 10 \cdot 0$ mm. (say), derived from them with the original polynomial values at $S = 10 \cdot 0$ mm. (table 6).
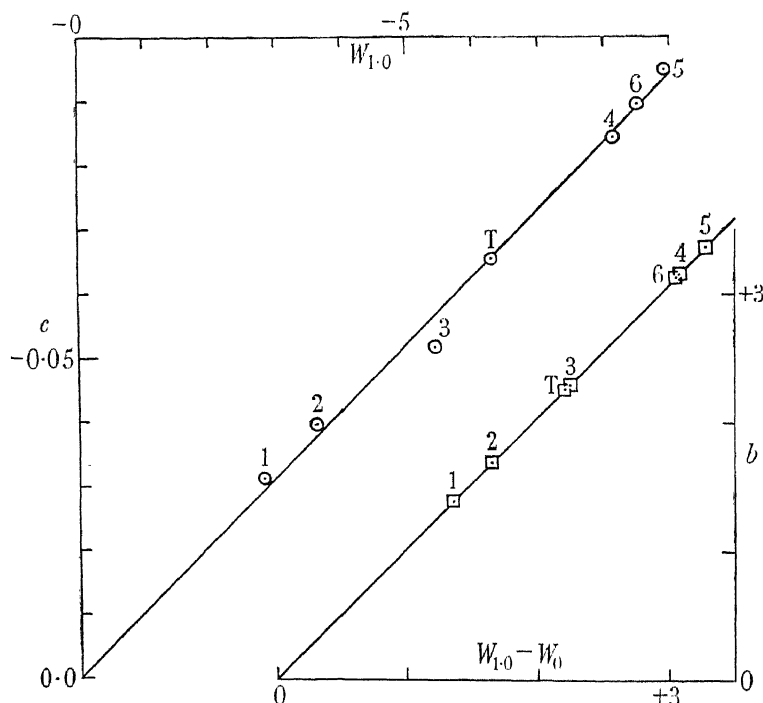
FIGURE 4. Graphs to demonstrate that the parameters a, b, c of the quadratic polynomials $W = a + bS + cS^2$, fitted to the relation between W and S, may be expressed as simple linear functions of two sets of standard width measurements, namely, W_0, the widths of the abdominal segments when $S = 0$, and $W_{1 \cdot 0}$, the widths when $S = 1 \cdot 0$ mm. a is itself the value of W_0, b is simply proportional to $(W_{1 \cdot 0} - W_0)$ and c is simply proportional to $W_{1 \cdot 0}$. The number of the segment is placed beside its point.

The quadratic regressions would seem (p. 127) to provide an adequate overall picture of relative growth for the range studied ($S = 4 \cdot 0$ to $S = 13 \cdot 25$ mm.). They yield a reasonably simple cinematogram, valid for the whole abdomen, though with two efficient standard width measurements, W_0 and $W_{1 \cdot 0}$, i.e. the widths at $S = 0 \cdot 0$ (table 1C) and $S = 1 \cdot 0$ mm. (table 4) respectively, as the irreducible minimum. As before a is identical with W_0. The other parameters b, c give elliptical curves when plotted against a (cf. p. 128), but (figure 4) c gives a very good linear relation, passing through the origin, to $W_{1 \cdot 0}$:

$$c = 0 \cdot 0105(W_{1 \cdot 0}),\tag{1}$$

and b gives an equally good linear relation, passing through the origin, to $(W_0 - W_{1 \cdot 0})$:

$$b = -1 \cdot 027(W_0 - W_{1 \cdot 0}).\tag{2}$$

The cinematogram may therefore be written

$$W_S = W_0 - 1 \cdot 027(W_0 - W_{1 \cdot 0}) \, S + 0 \cdot 0105(W_{1 \cdot 0}) \, S^2. \tag{3}$$

The adequacy of this key relation is tested in table 6A, where the values of $W_{10 \cdot 0}$ calculated from it are compared with the polynomial values at that size, $S = 10 \cdot 0$ mm.

The cinematogram, like the relations for the individual segments (table 1B), expresses the fact that relative growth rate is a continuously variable function of body size, i.e. that dW/dS is itself a function of S. The coefficient of S^2 in the cinematogram is negative in sign, since $W_{1 \cdot 0}$ is negative, and the change in relative growth rate is one of retardation (p. 122). Relative growth rate is thus a monotonically decreasing function (Medawar 1944, p. 139) of S. The change is not monotonic on cubic or quartic fitting, when a subsequent acceleration of relative growth rate is recognized (p. 124).

The change in W_S, at a fixed value of S, as a function of W_0, i.e. dW_S/dW_0, defines the change in width shape of the abdomen as a function of position (Medawar 1944, p. 139). The only variable in this expression, dW_S/dW_0 (derived from equation (3)), is $d(W_{1 \cdot 0})/d(W_0)$. This variable will be a complex expression since $W_{1 \cdot 0}$ is related to W_0 by the now-familiar elliptical curve, the definition of which will not be attempted here. Suffice to indicate that the change in shape, as a function of position, would appear to be determined entirely by the relation between the shapes at $S = 0 \cdot 0$ and $S = 1 \cdot 0$, all subsequent change being determined by that early change.

There remains the possibility of deriving the cinematogram by the alternative method, the direct method of Medawar (p. 116), the series of width measurements, W_S, at each value of S being in turn related to the series W_T, at a standard body size, as the first step, and the parameters of the set of relations so obtained then being related to body size as the second step. Here a familiar difficulty is encountered; the graph of W_S/W_T is typically an ellipse, and for values of S remote from T it is an even more complex loop, demanding an expression of quartic or higher degree for adequate definition (Ganguli 1931; Frost 1918; etc.). The extraction of shape only from the data does not simplify these curves (p. 116) but rather emphasizes their complexity.

A point of some interest, here, is that the relation between $W_{8 \cdot 0}$ and $W_{13 \cdot 25}$ is a very good straight line, passing through the origin, and the shape is therefore identical at these two sizes, but between $S = 8 \cdot 0$ and $S = 13 \cdot 25$ the shape changes considerably, the graphs of $W_S/W_{13 \cdot 25}$ becoming distinctly elliptical at intermediate sizes. On quadratic fitting this ellipsis is slight, and it is clearly increased by cubic fitting, which admits the upper inflexion on the W/S curves (p. 119). Below $S = 8 \cdot 0$ the distal segments, by comparison with the shape at $S = 13 \cdot 25$, are relatively narrower than the proximal group, and constitute the half of the ellipse lying on the abscissal side of the major axis ($W_{13 \cdot 25}$ as abscissa), whilst above $S = 8 \cdot 0$ proximal and distal groups 'change sides' of the major axis. It is as though the relatively high growth rate in distal segments, necessary to ensure appropriate adult proportions, leads to an overshooting of the mark, immediately above $S = 8 \cdot 0$, this being subsequently corrected so that relative proportions (i.e. the

shape) at $S = 13 \cdot 25$ become once more those at $S = 8 \cdot 0$. Here final proportions would appear to be more significant than the precise course of relative growth (cf. below).

The change in proportions between $S = 8 \cdot 0$ and $S = 13 \cdot 25$ is small by comparison with the changes below $S = 8 \cdot 0$ and, while the former might lead to a relatively simple cinematogram, the latter is scarcely amenable to the present methods unless, once more, the abdomen be divided into proximal and distal portions. It seems improbable, after various attempts, that this approach could lead to a simpler cinematogram than that already considered (equation (3)).

A remarkable simplification in the W_S/W_T curves may, however, be effected by 'excising' from the abdomen the portion already present at $S = 4 \cdot 0$, that is, by subtracting the width at that size from all subsequent values of segment width, and plotting $(W_S - W_{4 \cdot 0})$ against $(W_T - W_{4 \cdot 0})$. The graphs are now virtually straight lines, though not quite through the origin, and a trace of the elliptical curves is evident for values of S remote from both T and $4 \cdot 0$. The choice of $W_{4 \cdot 0}$ as the series of measurements to be subtracted is somewhat arbitrary, and it is possible that an even more useful series might be found. Empirical attempts to find this series, W_C, which would yield graphs $(W_S - W_C)/(W_T - W_C)$, all of which should be straight lines, passing through the origin, have not been wholly successful; while this approach deserves consideration, therefore, not only on account of the relative simplicity of the form change it implies but also on account of its possible biological significance (p. 127), it will not be developed further here, since it seems unlikely that the ultimate cinematogram could be a simpler expression than the one preferred in equation (3). The best approximation actually obtained is

$$(W_S - W_C) = a + b(W_T - W_C),\tag{4}$$

where the parameters a, b are both polynomials in S.

Since the inclusion of the initial width measurements of the abdomen so markedly complicates the picture of growth of the structure, over the size range studied, it seems reasonable to suggest that growth over this range is not a function of these initial measurements. It would appear to be related primarily to the observed growth gradient with its centre in segment 5 (p. 120, etc.), the quantitative details of the gradient being such as to produce functionally adequate final proportions and the initial measurements would appear to be virtually irrelevant. In this event the specific growth rate, that is, the growth rate per unit of existing tissue, which is normally the most useful measure of growth potential (Medawar 1945, p. 163), is here not merely without significance, but positively irrelevant.

A sharp change in growth pattern of the abdomen, in the region of $S = 4 \cdot 0$ mm., is clearly indicated (cf. p. 120). This is no doubt correlated with the transition from the juvenile stage to the adult, in which the abdomen of female Brachyura is greatly enlarged, in connexion with its incubatory function. The sharpness of the transition in *Pinnotheres* may be due to the enormous development of the female abdomen in the genus, and possibly also to the radical nature of the change in function of the structure. In view of the sharp break in growth pattern at $S = 4 \cdot 0$ mm. it is relevant to recall (pp. 120, 122) that the segmental values of W when $S = 0$, and

other characteristics of the growth relations, show a gradient along the abdomen which is entirely independent of the growth gradient for the period $S = 4$ to $S = 13.5$.

In a shape-cinematogram (p. 116) the absolute magnitudes of the form-cinematogram are replaced by relative proportions; in the present example all values of W_S would thus be divided by $W_{S\,max.}$, the value of W_S in the widest segment (or by some other selected standard width-measurement). The ratio $W_S/W_{S\,max.}$ would replace W_S on the left-hand side of the cinematogram. The shape-cinematogram, in fact, could be derived directly from the form-cinematogram (equation (3)) by dividing both sides of equation (3) by $W_{S\,max.}$ which is a polynomial expression in S (table 1). The shape-cinematogram derived from equation (3) would therefore be

$$W_{sh} = [W_0 - 1.027(W_0 - W_{1.0})\ S + 0.0105(W_{1.0})\ S^2]/(-12.20 + 3.37S - 0.095S^2), \quad (5)$$

where W_{sh} is the ratio $W_S/W_{S\,max.}$. The relation is not excessively unwieldy and probably compares favourably with the shape-cinematogram which would be derived from proportionate widths as the initial data.

III. *The co-ordinate transformation of the growing abdomen*

In order to visualize the changes in shape of the female abdomen Thompson's admirably pictorial method of co-ordinate transformation has been applied, in figure 5A, to the male abdomen, which differs little in shape from that of the female below $S = 4.0$ mm. The transformation shows clearly the centre of maximal growth rate in the mid-line, in segment 5, the graded decline in rate both proximalwards and distalwards, and the resulting position of the segment of maximum width in the mature female. Growth in length shows the same centre and gradients, but growth in length is much slower than width growth in any segment. This has been confirmed by measurements on a number of individuals, which also indicated that the graphs of segment-length/S are formally similar to those of W/S.

The transverse grid-lines become curved much less than the longitudinals, implying that length growth is more or less uniform in rate across the whole width of the abdomen, at any level. The uniformity is probably more exact than indicated by the figure, if allowance is made for the effect of projection of the somewhat curved female abdomen on to a plane surface. The growth centre of the abdomen would therefore appear to be a transverse band rather than a point in the mid-line.

In figure 5B, the reciprocal transformation of the mature female into the male (i.e. also into the juvenile female) abdomen is shown, since features not evident in the one transformation are sometimes revealed in the reciprocal.* In the present example the density of the grid-lines and their curvature around segment 5 possibly display more forcibly than the reverse transformation the high differential growth rate of that region.

---

* Undergraduates in this laboratory have shown that the transformation of the chimpanzee's skull into the human skull reveals a high differential growth rate of the frontal region as compared not only with the muzzle, but also with the parietal region, of a degree which could not be anticipated from the reverse transformation illustrated by Thompson (1942, p. 1082).
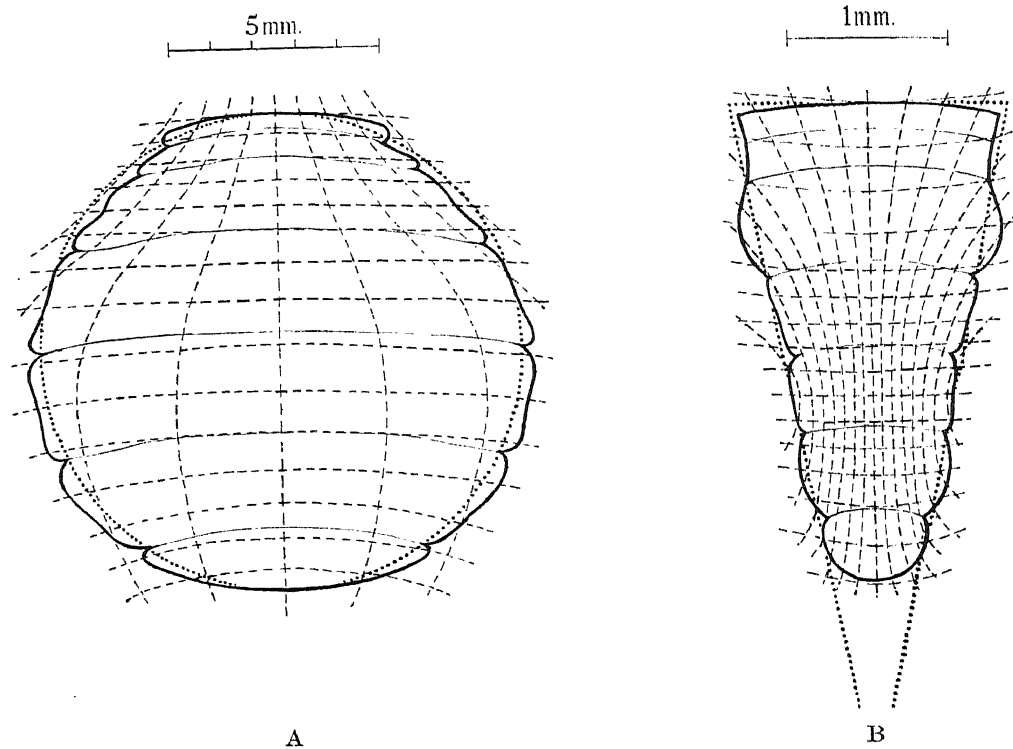
A. E. Needham

FIGURE 5. Reciprocal co-ordinate transformation between the abdomina of (B) the male (or the juvenile female) and (A) the adult female *Pinnotheres pisum*, reduced to the same overall length. Initially each abdomen was projected on to squared paper, and the grid shown in broken lines on each indicates how, in order to obtain that abdomen by transformation of the other, the square grid on the latter must be distorted. Growth centres and gradients are readily visualized. The dotted outline on each figure shows that the transformation is from a virtual triangle to a virtual circle.

From the dotted outlines superimposed on the two abdomina it is clear that the shape of the abdomen of the female changes, during growth, from a virtual triangle to a virtual circle.

CONCLUSIONS

As progressively higher powers of polynomial in carapace size (S) are fitted to the width (W) of the abdominal segments of this crab, progressively finer detail is revealed of the pattern of relative growth of the abdomen. The linear regression displays the main features (a) the centre of maximal growth rate in segment 5 and the gradient from this segment both proximalwards and distalwards in the abdomen, and (b) the progressively later time of onset of the main phase of growth, along the series of segments, in the proximo-distal direction. The quadratic regression is the best polynomial fitting the width of segments 1 and the telson, and it adequately defines the main phase of growth in the other segments. The best polynomial for the different segments is not necessarily strictly affine in all. The cubic regression suggests that the magnitude of the observed progressive retardation of relative growth rate is maximal farther forward in the abdomen (in segment 4)

than the segment (5) of maximal initial growth rate. The alternation of + and
— signs in the successive terms of the polynomial formally implies first a retarda-
tion, i.e. a negative acceleration, in relative growth rate, and secondly a check
on this negative acceleration, the negative acceleration therefore decreasing, at a
constant rate. The quartic regression gives expression to peculiarities of individual
segments which tend to obscure the general resemblances, and probably lies
beyond the point of maximal utility in the series of polynomials. However, the
values of S at the two points of inflexion on the W/S curves are more nearly con-
stant in all segments (omitting anomalous segments) than any characteristics of
the lower power polynomials.

The magnitude of the coefficients of the different powers of S in the polynomials
varies in a parallel manner along the series of segments, with a maximum in segment 4
or segment 5 and a gradient both proximalwards and distalwards. The gradients
become progressively steeper towards the two ends of the abdomen. Total and re-
sidual variance, and the variance due to each term of the polynomial, show the
same centre and gradients along the abdomen. The magnitude of the coefficients
varies in a similiar manner along the series of terms within each polynomial, being
maximal in the independent term and decreasing progressively in power of S. The
geometric rate of this decrease progressively increases in the same order, and the
arithmetic rate of this decrease progressively decreases in this order. The variance
due to the different terms shows the same gradient.

The set of quadratic regressions for the seven segments may be combined to give
an 'algebraical form-cinematogram' in abdomen width, by defining each of the
three parameters as continuous functions of abdomen width at two standard body
sizes. From this form-cinematogram the corresponding 'shape-cinematogram'
may be derived. The set of cubic regressions do not yield a single cinematogram
of reasonable simplicity, though two independent cinematograms may be fitted
to proximal and distal halves of the abdomen. This demands only one series of
standard width measurements.

A remarkable simplification of the form change of the abdomen during this
period is effected (p. 132) by excising, from all width measurements, the width
measurement of that segment at $S = 4.0$ mm., i.e. the width at the end of the
preceding phase of growth. This permits the derivation of a relatively simple
cinematogram by the direct application of Medawar's procedure.

The equation 'form = shape + size' would seem (p. 116) to be a useful convention.

REFERENCES

Atkins, D. 1926 *J. Mar. Biol. Ass. U.K.* 14, 475.
Fischer, A. 1946 *Biology of tissue cells.* Cambridge University Press.
Fisher, R. A. 1946 *Statistical methods for research workers.* Edinburgh: Oliver & Boyd.
Frost, P. 1918 *An elementary treatise on curve tracing.* London: Macmillan & Co. Ltd.
Ganguli, S. 1931 *Theory of plane curves*, 1. Calcutta: University Press.

Huxley, J. S. 1932 *Problems of relative growth.* London: Methuen & Co. Ltd.

Huxley, J. S. & Teissier, G. 1936 *Nature*, **137**, 780.

Lebour, M. V. 1928 *J. Mar. Biol. Ass. U.K.* **15**, 109.

Medawar, P. B. 1944 *Proc. Roy. Soc.* B, **132**, 133.

Medawar, P. B. 1945 'Size, shape and age.' In *Essays on growth and form*, ed. W. E. Le Gros Clark & P. B. Medawar, p. 157. Oxford: Clarendon Press.

Needham, A. E. 1937 *Proc. Zool. Soc. Lond.* A, **107**, 289.

Reeve, E. C. R. & Huxley, J. S. 1945 'Some problems in the study of allometric growth.' In *Essays on growth and form*, ed. W. E. Le Gros Clark & P. B. Medawar, p. 121. Oxford: Clarendon Press.

Sandoz, M. & Hopkins, S. H. 1947 *Biol. Bull. Woods Hole*, **93**, 250.

Thompson, D'A. W. 1942 *On growth and form.* Cambridge University Press.

Williams, G. & Needham, A. E. 1938 *Proc. Zool. Soc. Lond.* A, **108**, 539.

T

The work of the Discovery Committee

By N. A. MACKINTOSH, C.B.E., D.Sc., *Director of Research, Discovery Investigations*

(Lecture delivered 26 *January* 1950—*Received* 16 *February* 1950)

[Plates 8 to 11]

The geographical field in which most of the Discovery Committee's work has been carried out during the past 25 years is the Southern Ocean. This zone of continuous deep water, very rich in marine life, supports one major industry—the whaling industry—but is otherwise little developed as yet, and seldom visited. It is not easy to find a short descriptive label for the work itself, but nearly all of it comes under the headings of deep-sea oceanography, whales and whaling, or Antarctic geography, and much of it is concerned with the interrelations of these subjects.

Since the beginning in 1924 the Discovery Committee has worked under the Colonial Office, but in 1949 the Committee's functions, together with the scientific staff, the ships, and other assets, were taken over by the Admiralty, and now form part of the new National Institute of Oceanography. The Discovery Committee, in its original form, has been dissolved, but it is encouraging to know that the continuation of its work is assured.

ORIGIN AND PROGRAMME OF THE DISCOVERY COMMITTEE

The Committee was appointed in 1924 by the Secretary of State for the Colonies, to carry out the recommendations made in a *Report of the Interdepartmental Committee on Research and Development in the Dependencies of the Falkland Islands* (Cmd. 657, 1920). The waters of the Dependencies form a substantial sector of the Southern Ocean, and, at that time, constituted the world's most important whaling centre. The *Report* reviewed the natural resources of the region, and in view of the rapid development of the whaling industry, it was proposed that the investigations should bear largely on the biology of whales and on their oceanic environment, mainly with a view to the proper regulation of the industry; but it urged the importance of research on very broad lines and recommended the study of the hydrography and natural history of the whole area.

The Discovery Committee was an Executive Committee, acting on behalf of the Government of the Dependencies but responsible to the Secretary of State. The first chairman was Mr E. R. Darnley of the Colonial Office, and among members were representatives of the Admiralty, the Ministry of Agriculture and Fisheries, the British Museum (Natural History), and the Royal Geographical Society. Dr Stanley Kemp was appointed Director of Research, and the work was financed from a substantial 'Research and Development Fund', built up by the Government of the Falkland Islands from taxes on whale oil processed in the Dependencies. Captain Scott's former ship, the *Discovery* (figure 8, plate 10), was purchased and refitted for oceanographical research, a Marine Station was built at South Georgia

in the Dependencies (figure 4, plate 8), and in 1925 work in the field began as an expedition called the 'Discovery Expedition'.

The programme was to include research on the general biology of whales by examination of the carcasses brought into whaling stations, and on their distribution and movements by the attachment of marks to the living whale at sea. At the same time the factors governing the distribution of whales, and, indeed, the whole marine fauna and flora, were to be studied by means of a comprehensive oceanographical survey. Broadly these lines of approach are in accordance with the methods of modern fisheries research, but the expedition had to adapt them to a new object in a new field, and to develop certain new technical methods appropriate to long voyages in severe weather conditions, to intensive deep-sea observations, and to such large and inaccessible animals as whales. It had further been decided that the oceanographical work should not be confined to subjects of immediate economic application, and that general deep-sea research, in the tradition of the Challenger Expedition, should be included. This was an ambitious programme, and it needed a man of the calibre of the late Dr Kemp to put it into operation. The Discovery Committee was a strong committee which formulated a policy and guided the work, and in the first place, I believe, it owed its existence to the initiative of Mr Darnley; but the work of building up the Discovery Investigations devolved primarily upon Dr Kemp. He relinquished his post in 1936 to become Director of the Plymouth Laboratory but continued as a member of the Discovery Committee until his death in 1945, and at every step we are still conscious of the effects of his wise judgement and foresight.

The Committee's work in the first place was to launch an expedition, and thereafter there was little to be done in England until the first voyages of the ships were completed. It may therefore be best to give an account first of the work in the field, the methods employed, and the principal results, and to deal later with the organization which was built up in London and the arrangements for research at home.

THE 'DISCOVERY EXPEDITION', 1925–1927

The Dependencies of the Falkland Islands lie between the meridians of 20° and 80° W. (figure 1), and they include part of the Antarctic continent and several scattered islands and groups of islands separated by a few hundred miles of deep ocean. Most of them are outside the Antarctic circle, but they are all mountainous and heavily glaciated. Whaling was carried out from land stations at South Georgia and moored factory ships in the South Shetland Islands and South Orkney Islands, but it has now ceased at these shore bases except at three of the stations at South Georgia.

The Marine Station with its laboratory was built close to one of the whaling stations at South Georgia (plate 9), and work began here early in 1925. Each whale- brought into the station was measured and examined, and we were interested mainly in their breeding, growth, and age, their food, and the range of variation in their dimensions and external characters. The observations made were partly new, and partly based on methods initiated by other biologists, but the principal object was

to get data from a sufficient number of whales for quantitative treatment. Good progress was made, and by April 1927 we had examined nearly 1700 whales. With some changes of staff the station was kept open until 1931, by which time we had

FIGURE 1

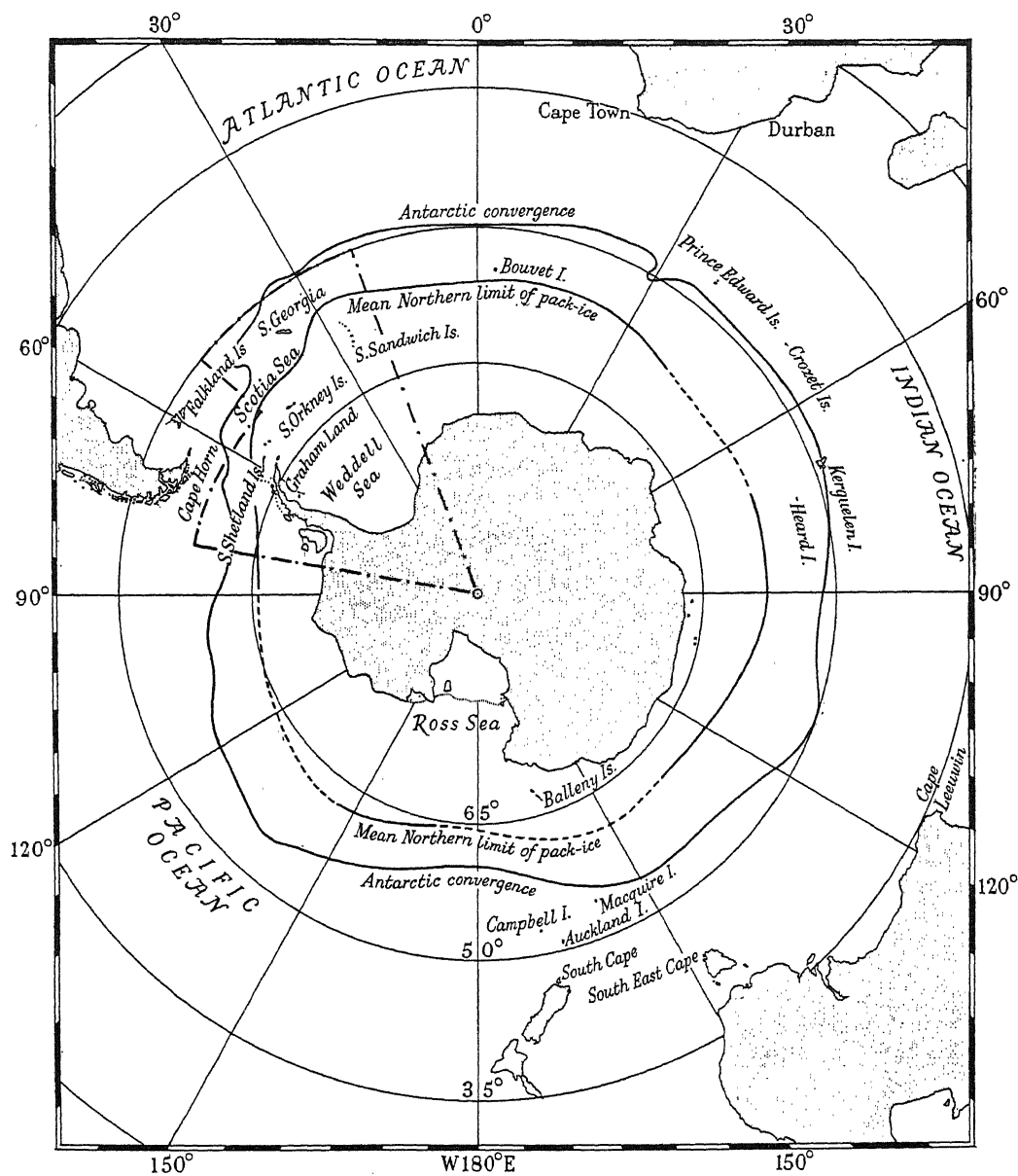the 'dossiers' of about 3700 whales. I shall return to the results of all this work later. In the meantime the Royal Research Ship *Discovery* undertook a preliminary survey of the South Georgia whaling 'grounds' in 1926 and was later joined by the R.R.S. *William Scoresby*. More thorough ocean surveys of the whaling grounds were then carried out around South Georgia and the South Shetland Islands, and

lines of deep-sea 'stations' were extended over the Falklands sector of the Southern Ocean and parts of the South Atlantic. In 1927 the *Discovery* returned to England, but in the next two years much of this work was repeated and extended by the *William Scoresby*. It can be looked upon now mainly as a preliminary, but intensive, survey of one part of the Southern Ocean, and it revealed the general content and distribution of the plankton population, and the local distribution of water masses and currents.

THE *DISCOVERY II* AND EXPANSION OF THE FIELD

The Committee had not been in a position at first to plan a scheme of research extending over any long period, and in 1927 the future of the work was still very uncertain. In the following year, however, the *Discovery* was taken over for the British, Australian and New Zealand Antarctic Expedition, and it was then decided that a steel steamship, the R.R.S. *Discovery II*, should be built, with greater range and power, and specially designed for deep-sea research.

It was the acquisition of this ship in 1929 which converted an expedition into a long-term service. It changed the whole complexion of the work, and the title 'Discovery Investigations' became more in keeping with its nature. For a proper understanding of the Discovery Committee's work I think it is essential to regard it as something intermediate between an expedition and a permanent research institution. In any new venture results are achieved, at least in part, through trial and error, and although an expedition may make new discoveries of great importance, it is often a matter for regret that when it is over, hard-won experience is lost through the dispersal of the staff to other occupations, and valuable equipment, often capable of long-continued use, is disposed of. The Discovery Investigations have evolved in some ways perhaps in a haphazard, hand-to-mouth manner, with some of the adventures, mistakes and new achievements of an expedition; but the strength of the present organization lies in the fact that after breaking new ground it has been able to accumulate experience and retain its capital equipment, and that it can count on the help and co-operation of many scientists who have been associated with it.

After the *Discovery II* was commissioned the field of work expanded. The preliminary work of the old *Discovery* and the *William Scoresby* had shown that although local oceanographical conditions could be measured and mapped, they could not be properly understood without fuller knowledge of the water masses and currents of the whole circle of the Southern Ocean. Since whales travel great distances they also cannot be adequately studied in a restricted area; and furthermore, with the new pelagic factory ships the whaling industry was now spreading round most of the Antarctic seas. The work of the *Discovery II* eventually became in effect a survey of the whole Southern Ocean—and this, incidentally, was a task far beyond the capacity of the old *Discovery*. It needed a ship with greater power and cruising range; but since this ocean forms a circumpolar belt, not divided by land masses, the water circulation is relatively straightforward, and in consequence some at least of the problems are simplified. It is worth noting here that the *Challenger* and other expeditions have given us a broad picture of the oceans as

a whole, and if the work of the *Discovery II* had been planned purely as a further step in the exploration of the oceans—that is to say, with more intensive observations in one ocean—the Southern Ocean might well have been chosen as the region most favourable for the basic study of oceanic phenomena on a large scale.

METHODS AND EQUIPMENT

Ocean surveys of the kind undertaken by the *Discovery II* consist essentially in making observations at suitably scattered points, and interpolating between those points. At intervals varying between about 10 and 200 miles the ship is stopped on 'station', a sounding is taken, and then, with special sampling bottles and deep-sea thermometers, water samples and temperature readings are taken at various levels from the surface to the bottom. Salinity, oxygen content, pH and certain nutrient salts are estimated in the water samples, and density is calculated from the temperature and salinity. At the same time the plankton is collected with vertical and oblique closing nets at different depths, and these sample everything from the shrimp-like 'krill' (*Euphausia superba*), which constitutes the food of whales and many other animals, down to the minute constituents of the phytoplankton (mainly diatoms), which form a basic supply of vegetable food, comparable in a way to the grass on land. From a line of stations these physical, chemical and organic constituents can be mapped in vertical sections of the ocean; a network of stations gives their distribution in three dimensions; and lines repeated at different times of year reveal the seasonal variations. It is a common experience in work of this kind that the more crowded are the observations, the more complex are the conditions found to be. The area of the Southern Ocean is of the order of 30 million square miles, and here, of course, we are dealing with very widely scattered observing points, so that isolated observations, or groups of them, may not be of much significance. However, when the data are studied as a whole an unmistakable pattern does emerge, and certain valid conclusions can be drawn as to the general distribution and movements of the principal water masses, and the way in which they control the distribution of the plankton populations in general and the food of whales, and thus affect the whales themselves.

In deep-sea research we are groping in an invisible world with instruments which can at best give incomplete knowledge, and progress in such investigations is very largely dependent upon, and limited by, progress in the invention of technical devices. Indeed, any new device—echo sounding is an example—will open up new avenues of research, and will often discover phenomena which were not previously known to exist. Even with the methods already available it is obvious enough that for a survey of such a vast area as the Southern Ocean the whole subject of the design and management of ships, and the supply, operation and maintenance of the scientific equipment must loom very large in the organization of the work. Technical problems are greatly increased when oceanic depths are to be explored, and in fact the stage at which any material is examined in the laboratory follows a long series of administrative and practical operations. In this connexion it is significant that the Committee's scientific staff, which usually numbered about eighteen to twenty,

must be supported by an executive staff in the ships of about seventy-five, and a secretarial staff of four or five. I shall explain later how a good deal of the scientific work is farmed out to specialists outside the Committee's organization, but the secretariat also would need to be much larger if we did not have liberal assistance from the common services of a large government department. The management of deep-sea research ships operating far from home involves heavy administrative work, especially before a ship sails on a new commission, but much of this, including arrangements for docking and repairs, placing orders, appointing agents abroad, paying accounts and emoluments, and so forth, has been executed by the Crown Agents for the Colonies, and in our new organization under the Admiralty corresponding services are available.

The *Discovery II* (1036 tons gross) is larger than the average research ship. She is an oil-burning steam ship, with some protection against ice, and has a maximum cruising range of 10,000 miles. The principal items of equipment are the light steam deck engines with reels carrying up to 3500 fathoms of 4 mm. wire for water sampling bottles and vertical nets, and a heavy winch with a small drum for 1000 fathoms of warp, and a large drum for 5000 fathoms of tapered warp. These are for the larger tow nets, trawls, dredges, etc. The instruments and accessories which are used, the sounding machines, the laboratory equipment, etc., could hardly be described here, and it is enough to say that all normal deep-sea operations are provided for, and spares are carried to allow for prolonged and intensive work. The old *Discovery* (736 tons gross) was a wooden auxiliary barque, and her equipment was similar to, though less comprehensive than, that of the *Discovery II*. The *William Scoresby* (324 tons gross) carries a commercial otter trawl and equipment for limited oceanographical work, and she incorporates some of the features of a trawler and some of a whale catcher.

OCEAN SURVEYS BY THE *DISCOVERY II*, 1929–1939

Each commission of the *Discovery II* lasted for about 20 months. In the first of these (1929–31), which was led by Dr Kemp, she worked in the Atlantic sector of the Southern Ocean, but the second commission under Mr John included a complete circumpolar cruise in the winter months, with long V- or W-shaped lines of stations between the southern continents and the ice-edge. The third, fourth and fifth commissions (1933–9), which were led by myself, Dr Deacon and Dr Herdman respectively, included repeated lines of stations on the Greenwich meridian and in 80° W (South-east Pacific), cruises tacking across the summer zone of whales and krill north of the pack-ice, a further circumpolar cruise in summer, and a variety of shorter cruises for subsidiary purposes. Figure 2 shows the ship's principal tracks (omitting some intensive work in the Falklands sector and coastal surveys), and the lines can mostly be taken to represent rows of stations and almost continuous soundings.

The picture of the Southern Ocean which emerges from this work is a series of circumpolar zones or gradients, distorted by the effects of land masses and bottom topography, but recognizable in all longitudes. Any line of stations running from north to south will cut across certain distinct winter masses and several zones with

distinct populations. But from east to west there is little change, and the major features of the Southern Ocean can be seen in a vertical section in almost any meridian.

It was already known that the general drift in this ocean is from west to east (except for a counter current close to the Antarctic coast), and the German research ship *Meteor* in 1926 showed that in the Atlantic sector there is an Antarctic 'con-
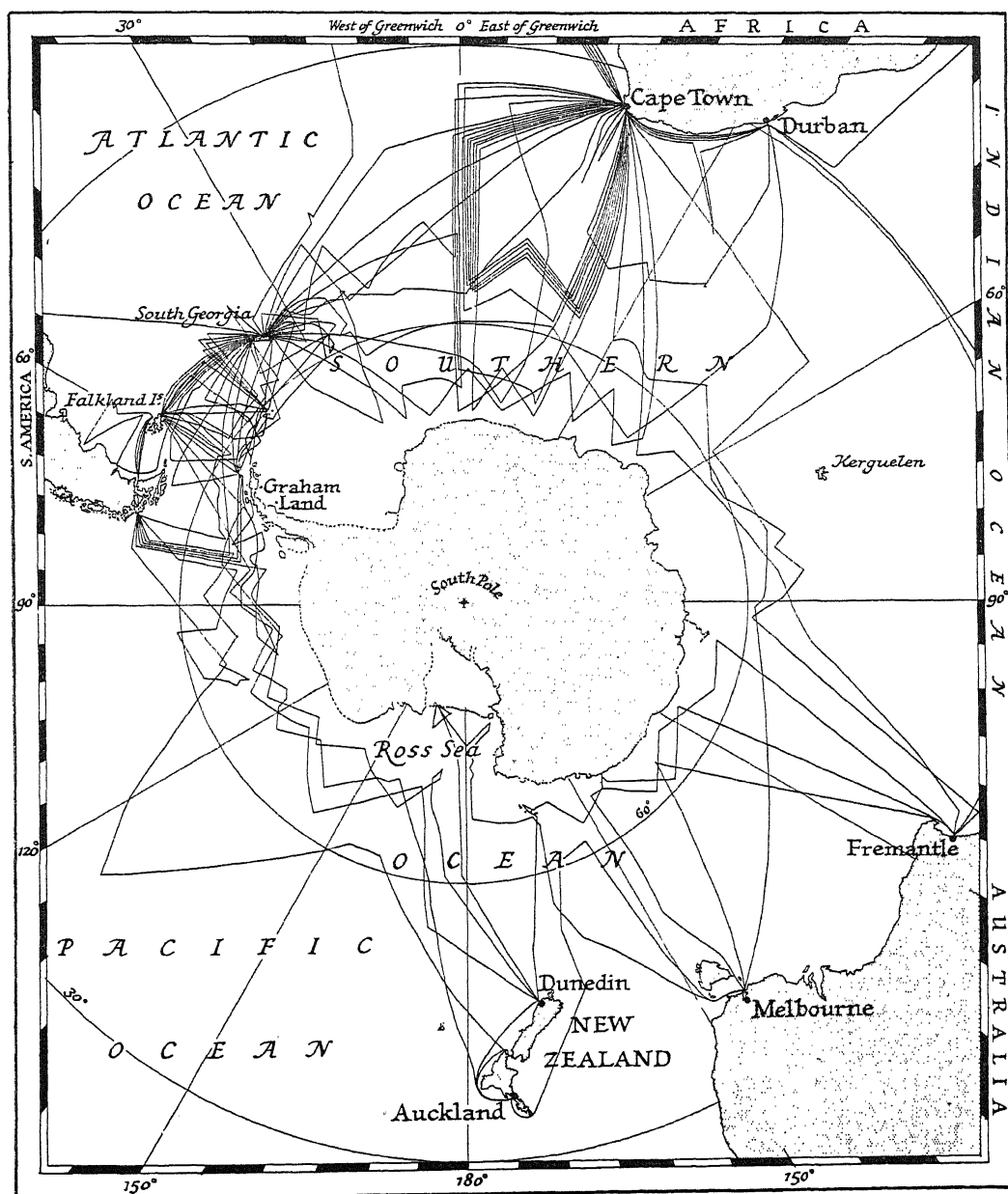

FIGURE 2. Principal voyages of the *Discovery II*, 1930–9. (The tracks are only approximately correct. Where they are crowded together they have been straightened and separated.)

vergence' which marks the northern limit of the cold Antarctic surface layer (see figure 1). In the eastward movement of this layer there is a component towards the north, and the Antarctic convergence is the point at which it sinks below the sub-Antarctic water to form the 'Antarctic intermediate layer' (which can be traced northward beyond the equator). Below the cold surface layer is a warmer, more saline layer of great depth, with a compensating southerly component (figure 3); and below that is the Antarctic bottom water, again with a northerly component. I am inclined to think that the most important single fact demonstrated in all the work of the *Discovery II* is that these three layers and the Antarctic convergence extend all the way round the Southern Ocean. That is quite a simple fact, and it would be surprising if it were not so, but it is all important that such comprehensive features should be established beyond any doubt.
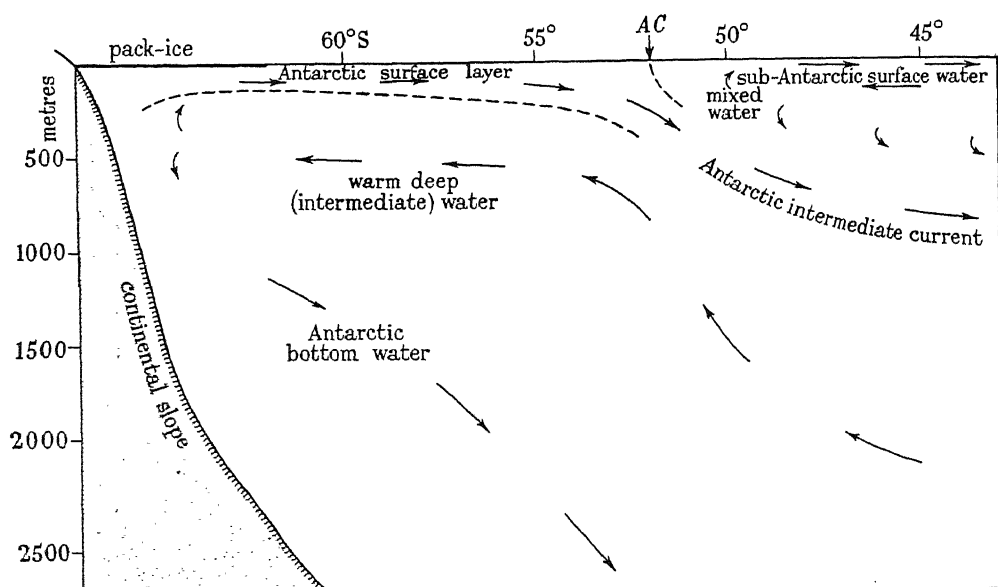
FIGURE 3. Diagrammatic section of the Southern Ocean, showing north and south components in the movements of the principal water masses in high southern latitudes. AC = Antarctic convergences. The latitudes of the convergence, pack-ice edge, and the continental coast are variable.

Most of the work of the *Discovery II* has consisted in exploring the variations in this basic pattern, and studying its relation to the pelagic fauna and flora. This is not a matter of just mapping out and polishing up the details; we are studying the anatomy and physiology, as it were, of the ocean. There are the seasonal changes of far-reaching significance, especially to the fauna and flora, and there are the permanent effects of topography. The Weddell Sea, for example, is a huge bay in the Antarctic continent, and its cold deep water is the principal source of the heavy Antarctic bottom water, which creeps east and north right round the Antarctic. Its northward movement appears to determine the latitude in which the warmer and more saline intermediate water is deflected upwards; and where this warm layer slopes upwards to the south the relatively light Antarctic layer above it can slide

FIGURE 4. The Marine Station at Grytviken, South Georgia.

FIGURE 5. The Laboratory at the Marine Station.

FIGURE 6. Grytviken, South Georgia. View of the harbour.

FIGURE 7. A blue whale on the flensing platform at Grytviken.

FIGURE 8. The R.R.S. *Discovery* at Port Lockroy, Palmer Archipelago.
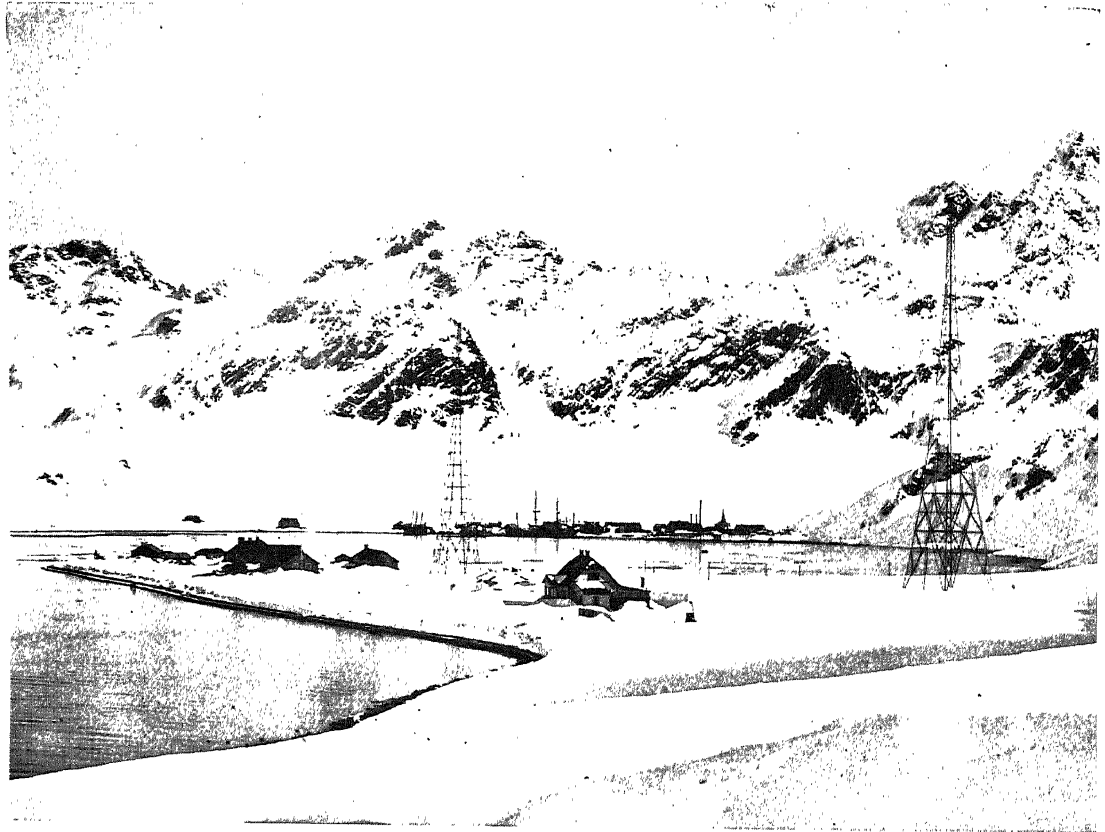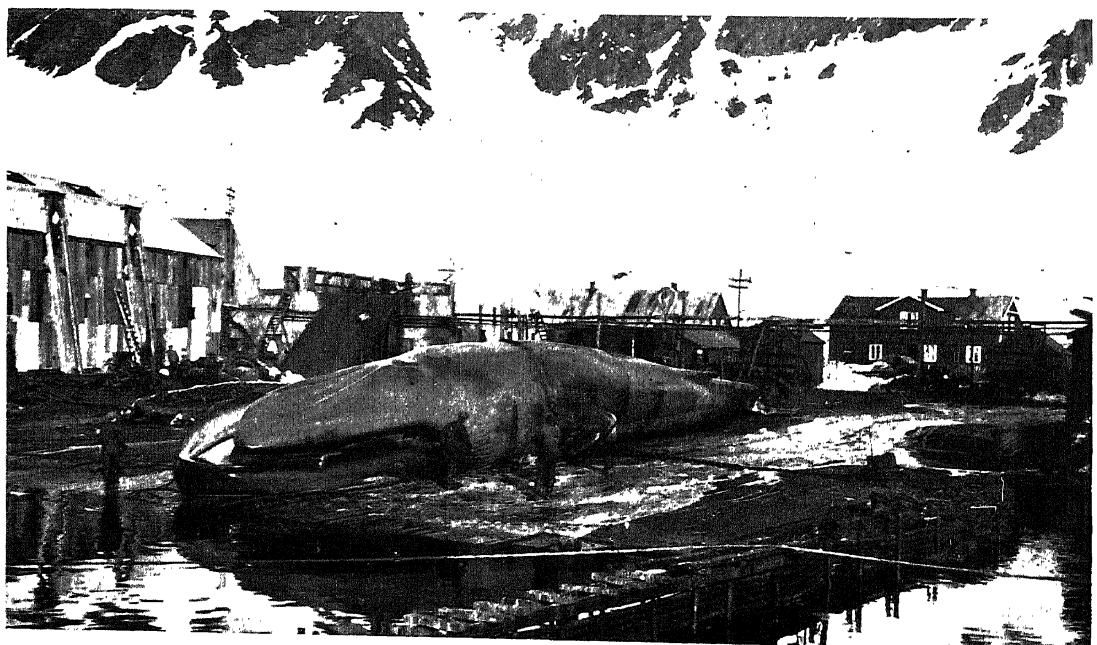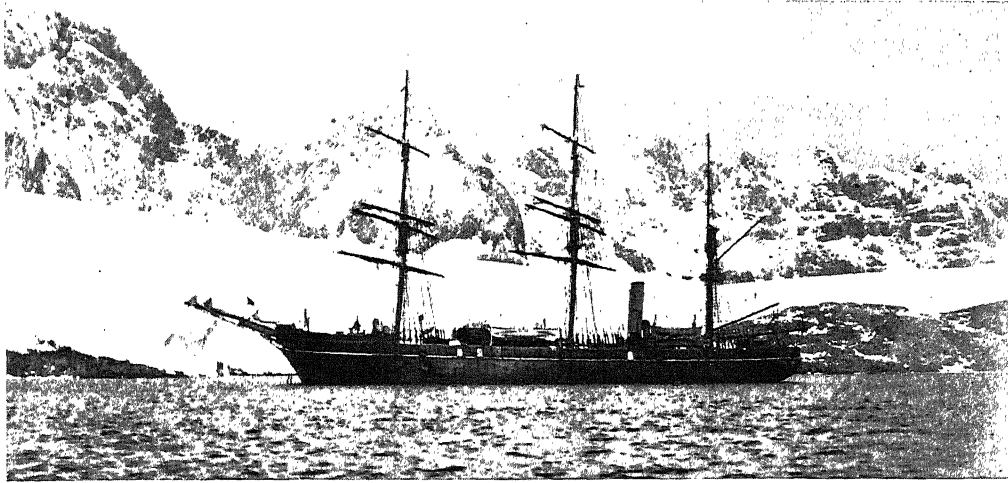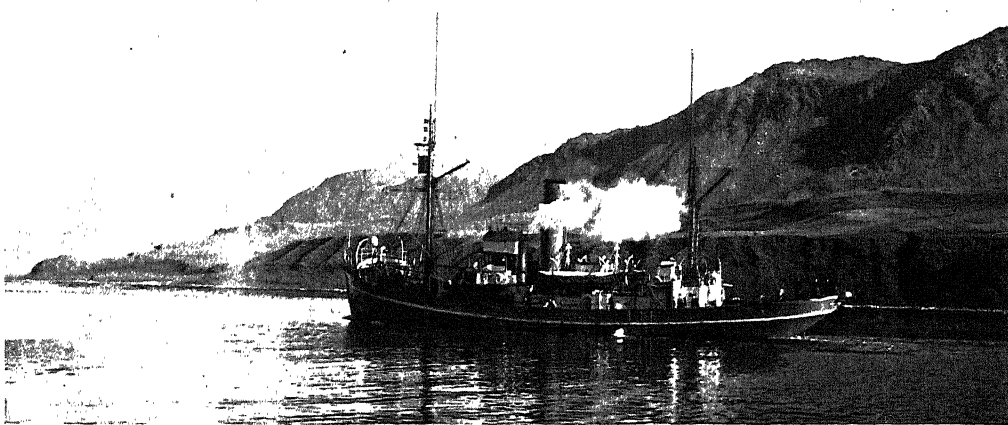
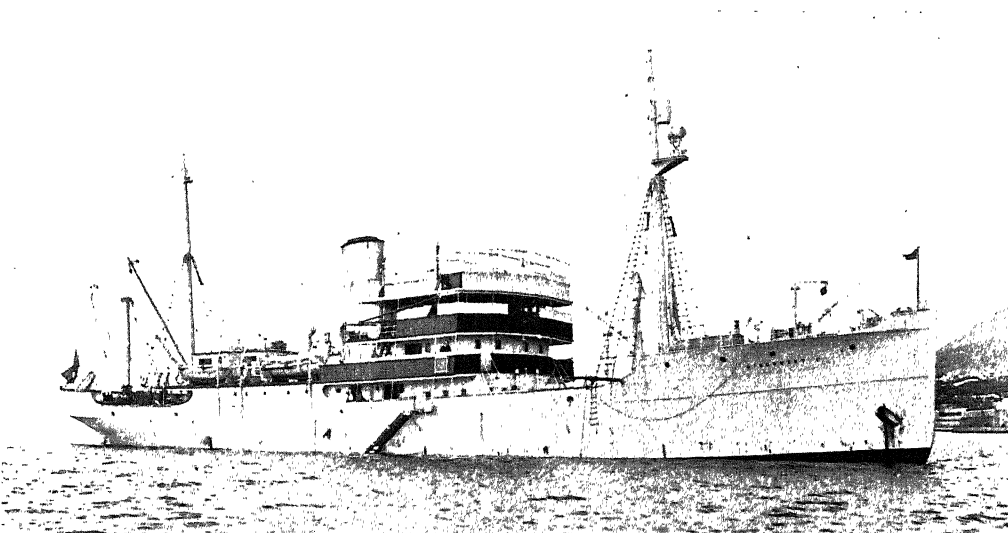FIGURE 9. The R.R.S. *William Scoresby* at South Georgia.

FIGURE 10. The R.R.S. *Discovery II*.

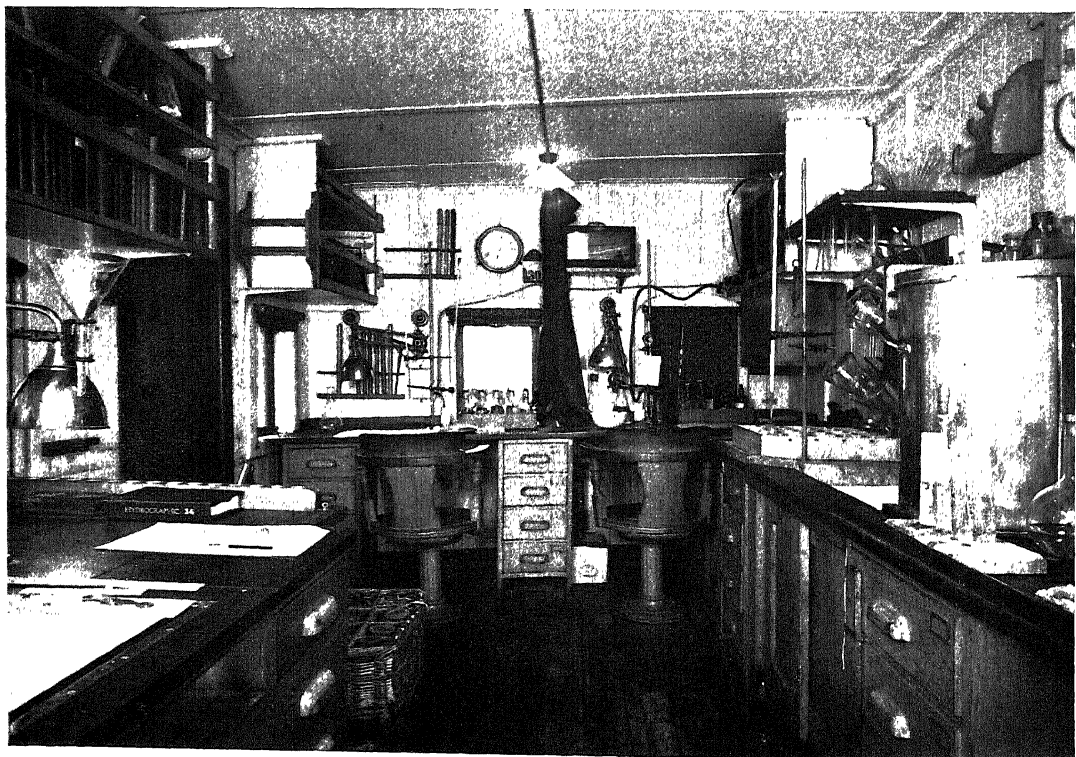FIGURE 11. Biological laboratory in the *Discovery II*.

FIGURE 12. Chemical laboratory in the *Discovery II*.

downwards to the north, underneath the sub-Antarctic surface water. Hence it is inferred that the position of the Antarctic convergence at the surface is determined by relative movements of the intermediate and bottom currents.

In the upper part of the Antarctic surface layer there is a rich growth of diatoms and animal plankton. Organic matter sinks and enriches the deep intermediate water with nutrient salts which are carried south and back to the surface. The key to the distribution of pelagic organisms in the Antarctic, including the food of whales, has been found in the relative movements of the surface and deep intermediate layers. We have found that the various species of the animal plankton have different devices for spending part of their lives moving east and north in the surface layer, and part moving east and south in the deeper water. It can scarcely be doubted that this is the mechanism by which they maintain the boundaries of their normal distribution—some by diurnal migrations from one layer to the other, many by spending the summer in the surface and the winter at a deeper level, and some spending different parts of their life cycle in the two layers. Apart from the shoaling krill (*Euphausia superba*), those species which form the bulk at least of the macro-plankton exhibit the second of these methods. Concentrated in the surface layer in summer they almost desert it in winter, having sunk far into the southward-moving intermediate water; and both the sinking in autumn and the upward movement in spring take place earlier in the lower than in the higher latitudes. This seasonal migration must set up a large-scale circulation which brings about the replenishment of the stocks in the far south. The krill gives an interesting example of the third method. The adults live very close to the surface, but their eggs sink to 1500 m. or more, and the successive larval stages are passed between increasingly shallow horizons as they migrate through the intermediate layer towards the surface. It must be supposed that in the general eastward drift, each species, in a succession of generations, will circulate perpetually around the Southern Ocean, and that by adjustment of its movements between the two layers it can keep within a more or less restricted zone with a suitable range of temperature and perhaps other conditions. This may require final confirmation, or it may be an over-simplification of what actually takes place, but it does seem to offer an adequate explanation of the principal features of the plankton distribution.

DIRECT RESEARCH ON WHALES

Turning now to the more direct investigation on whales I should say something of the method of marking them. This was an entirely new undertaking, but after considerable experiment had been made it was found that a stainless steel tube fired from a shot gun was successful. The marks are designed to bury themselves under the blubber, and each has a serial number and an inscription offering a reward for its recovery by the whalers. About 5000 whales have been marked in various parts of the Southern Ocean, mostly by the *William Scoresby*, but also by hired whale catchers near South Georgia, and so far nearly 300 marks have been recovered. (An unknown proportion may be lost through a healing process.) These recoveries have had several important results. They have clearly proved that at least some whales do (as was believed) undertake long migrations between the Antarctic and tropical

waters; they have shown that a whale usually, though not always, returns after its winter migration to the same part of the Antarctic in the following summer; and they have shown that the species differ in their tendencies to become segregated into distinct communities. The proportions in which the marks are recovered from the different species are instructive; and as time goes on the marks will no doubt provide information on the ages and length of life of whales. Some marks recovered recently had been carried by whales for periods up to 14 years.

The anatomical examination of whales at South Georgia was extended to South African whaling stations for two winter seasons, and more recently to the modern pelagic factory ships, and altogether some 6000 whales have been measured and examined. The majority of these were fin and blue whales, but humpback, sperm, sei, and a few right whales were included. The blue whale (*Balaenoptera musculus*, figure 7, plate 9) is the largest species, with a maximum length of about 100 ft., and the fin whale (*B. physalus*)—not very much smaller—is the most plentiful. These two, and in some regions the humpback and sperm whales, are the mainstay of the whaling industry. From this work a large body of statistical data has been obtained on the range of variation within the species, and information has been gathered on their food, parasites, etc., but attention has been paid more especially to their breeding, growth and age, and to the constitution of the local populations of whales. The breeding season, period of gestation, rate of breeding, and average lengths at birth, sexual maturity and full growth have been approximately established, and provisional estimates of the rate of growth have been made. As a general rule they appear to grow fast but to breed rather slowly. Methods have been found also for judging the relative, though not yet the absolute, age of individual whales.

The method of age determination will serve as an example of the technique of work in whaling factories. The best criterion we have found is the number of old corpora lutea, or corpora albicantia, in the ovaries of an adult female.* Each of these corpora lutea records the shedding of an egg from the ovaries, and there is good evidence that they persist for many years if not through the life of the whale. The number of them in a pair of ovaries can be correlated with the attainment of full growth (which can be checked by examination of the vertebral epiphyses), and so we can take this number as at least a rough indication of the relative age of a female whale. We cannot yet translate numbers of old copora lutea with certainty into years of age, but the average rate of accumulation is probably about one a year. The physical work involves excavating the genitalia from a huge mass of viscera, and cutting notches in the vertebrae with an axe.

The work in whaling factories is carried out in unattractive conditions, but it does not require the elaborate equipment and technical experience needed for the deep-sea oceanographical work. The reason is of course that we have not had to catch our whales. This has already been done for us by the whalers, and I must acknowledge our great indebtedness to the whaling companies (too many to name here) who have not only given us many facilities for our work, but have regularly collected specimens for us over a series of years.

---

* Professor Ruud, of Oslo, has recently found a method of estimating the age of young whales by examination of the baleen plates.

From the data collected in whaling factories we could learn much more than we have learned about the constitution of the populations of whales if the catch were a fair sample of the whales in the sea. The whalers take whales where they can find them, and in a broad sense the catch is taken at random. But to some extent it must be biased by varying economic and political factors, weather conditions, and the whaling regulations, and it is often difficult to know how much, and what form of, selection may result. For example, there was a gradual increase over a series of years in the percentage of pregnant blue whales in the Antarctic catches. It was difficult to find any explanation of this in changes in the areas, methods, or times of hunting, and it almost seemed that the rate of breeding was actually increasing; and yet the explanation might lie in some small and gradual change in the method of hunting—perhaps a tendency for the factories to work, say, a little farther away from the pack-ice edge, and so to sample a slightly differently constituted population. However, there are conclusions about the populations which can be drawn with some confidence, and we have good evidence on such matters as the relative abundance of the different species in the Antarctic population, of the sex ratio in baleen whales, and of certain changes in the make-up of the population during the summer season.

The Discovery Committee's work on whales has generally helped to give a solid foundation for the regulation of whaling, and many of the conclusions reached have specifically affected the regulations or assisted in the interpretation of the statistics of catches. For instance, simple estimates of the average length at sexual maturity have contributed to the protection of immature whales, and allowed the percentage of immature whales in the whole catch to be calculated from year to year. Observations by both ships, and the recoveries of whale marks, have provided evidence not only that blue whales are much scarcer than fin whales, but that a larger proportion of the stock of blue than of fin whales is killed. Counts of old corpora lutea have shown significant changes in the age distribution of the catches of blue whales, and it is partly in consequence of such findings as these that the protection of this species has been strengthened. Directly or indirectly the work has further contributed to the delineation of sanctuaries, an overall limit to the total catch, measures for the protection also of humpbacks, and the limits to the open season. The international regulations are not of course based only on the Discovery Investigations. The Norwegian scientists for instance have made important contributions, especially in the analyses of the statistics of catches.

OCEANOGRAPHY AND THE DISTRIBUTION OF WHALES

Although the survey of the Southern Ocean has become, in its own right, one of the principal aspects of the whole programme, the oceanographical work is an essential part of the research on whales. Without a knowledge of the oceans in which the whales live we could not see our problems in their right proportions, but more specifically this knowledge is concerned with their distribution, movements and habits. The majority of whales spend the summer in polar waters where they concentrate on the rich feeding grounds, and in winter there is a migration to lower

latitudes where breeding takes place. It is assumed that they seek warmer water for breeding, and that finding little to eat, they depend for their energy on the fat stored in their blubber and other tissues. In the summer their distribution must depend largely on their food, though they are directly influenced also by such factors as the position of the pack-ice and almost certainly the sea temperature. Cause and effect are to be examined, and it is important to have one's objects clearly in mind. The ultimate aim, I should say, is prediction. We should like to be able to say where, when, and in what quantities whales will be found, what will be the effect of hunting them at one time and place or another, and so on. For this we must try to relate the pattern of distribution of the plankton and whales on the one hand to the more readily defined physical and chemical pattern of the sea on the other, and to connect the fluctuations in their occurrence with fluctuations in such conditions as the ocean temperatures and currents. We should even try to go further back to the primary factors—the climate, and the winds and other variable weather conditions which affect the sea. This is a long road to travel, and the physics and chemistry of the Southern Ocean still need further study. However, some definite progress has been made. Knowledge of the plankton circulation is a step forward, and we are in a position now to draw fairly reliable maps of the distribution and density of the krill and the phytoplankton in the Southern Ocean, and of the principal physical and chemical conditions on which their distribution depends. Indeed, we could now, up to a point, define in oceanographical terms the circumpolar zone which whales seek in the summer months. This can be better done when certain gaps in the general survey are filled, but in many ways we can go further than this. Some years ago, for instance, Professor Hardy showed how the local distribution of whales could be predicted from the distribution of phosphates in the water. A number of other correlations have been explored, and although there is still much to be done I think we shall be able to link up more and more the oceanographical work and the direct observations on whales.

SPECIAL AND SUBSIDIARY INVESTIGATIONS

The Discovery Investigations have been focused very largely on the biology, or perhaps I should say the bionomics, of whales, and the ocean surveys might never have been undertaken were it not for their bearing on the organic resources of the Antarctic seas. Although there has been no departure from the original objects, the work has in many ways spread far beyond them, and its ramifications have extended into many other fields. This I think commonly happens when a problem is approached on a broad basis, and certainly where well-equipped research ships are concerned. The principal work of the *Discovery II* has been to survey the major features of the Southern Ocean; this in itself is a step in the general exploration of the sea, but it is a type of survey which can also take in its stride many special and subsidiary investigations. Much new information, for example, has accrued on the distribution and seasonal changes in the Antarctic sea ice, innumerable deep-sea soundings have revealed many varieties of bottom configuration; large collections of the marine fauna and flora have been added to the national collections at home;

the biology of seals has been studied; coastal surveys and soundings have improved the Admiralty charts; and the latest edition of the *Antarctic Pilot* is full of contributions from the officers of the research ships. Material obtained for others to work on includes routine observations for the Meteorological Office, and the collection of specimens for teaching departments and individual research workers. The *William Scoresby* also, in addition to her principal work of marking whales and assistance in general oceanography, has undertaken two major operations: an exploratory trawling survey, divided into three periods, of about 150,000 square miles of the Patagonian continental shelf, and an oceanographical survey of the Peru coastal current. Marketable fish were found during the trawling, but not in sufficient numbers to offer very good prospects of a commercial fishery. The value of the Peru current survey lay mainly in the accurate delineation of the current and in the measurement of its physical features, especially of the subsurface layers about which little was previously known.

ADMINISTRATION AND RESEARCH AT HOME

In giving a brief account now of the Committee's organization at home I must speak in the past tense, for although the scientific work both at home and in the field is continuing on much the same lines as before, the administrative system under the Admiralty is of course not quite the same as it was under the Colonial Office. The Discovery Committee was constituted as a team which included members qualified to advise not only on purely scientific matters but also on such subjects as the whaling industry, navigation and the management of ships, the Antarctic regions and organization of expeditions, and finance; and each member played an active part in the Committee's work which sometimes involved rather more than attendance at meetings. From 1937 onwards the meetings were also attended by observers representing the governments of Canada, Australia, New Zealand and South Africa. Since the scientific programmes broke new ground in many directions as time went on, and the organization of what might be regarded as a series of expeditions involved much administrative work, it was found convenient to hold rather frequent meetings of the main Committee (about nine times a year) to decide matters of principle and to receive the recommendations of a Scientific Subcommittee and sometimes of a Ship Subcommittee.

The headquarters of the Discovery Investigations have always been in London, and although we have not possessed premises in which all the administrative and scientific work could be brought together, a system was evolved which worked very smoothly. Committee meetings were held in the Colonial Office, the scientific staff were accommodated in the British Museum (Natural History), and the Director and Secretary had offices in Queen Anne's Chambers, in Westminster, within very easy reach of the Museum on the one hand, and the Colonial Office, Admiralty, Fisheries Department, etc., on the other.

The results of the Committee's work are published in the *Discovery Reports*, which are printed by the Cambridge University Press. There has been a more or less steady output of these reports since 1929, and so far twenty-five volumes have been issued.

At sea, or elsewhere in the field, the original data receive some preliminary treatment, but this of course must be followed up by prolonged study of the material at home before papers are ready for publication. A ship like the *Discovery II* can accumulate data at an embarrassing rate, but although there is still a considerable mass of material which has not been examined, we have been able to keep pace with the more important work and to deal gradually with the less urgent aspects of it. This however could not have been done without the generous co-operation of other bodies and persons, and we are specially indebted to the Trustees and staff of the British Museum (Natural History) who have provided accommodation for our staff and collections, and given constant advice and help in our work.

Most of the research on the biology of whales, and on the plankton and physical oceanography of the Southern Ocean—that is to say in the main stream of the Discovery Investigations—has been carried out by members of the Committee's scientific staff (or sometimes by former members). The duties of the staff have varied from time to time, and each member has had experience of several branches of the work. Of about fourteen scientific officers in the years before the war the majority were zoologists and three were chemists; and most of their time in London was free for research, though there was also some work to be done in the supervision of the collections, library and scientific equipment and occasional assistance in the Director's office. The *Discovery II* generally carried two or three zoologists, and one or two chemists, with three assistants and certain members of the crew detailed for scientific duties. In the *William Scoresby* there might be one or two zoologists, and others might be working in whaling factories or on other detached missions. At any one time, however, about two-thirds of the staff would be working in London, or more when the *Discovery II* was home. Naturally we do not attempt to work on the whole of the data from one voyage before dealing with material from the next voyage; the method is for any one worker to take up some particular subject at a suitable stage (say the seasonal periodicity in the phytoplankton, or the bottom topography of some region) and to work through all the relevant data available up to date. The choice of such subjects depends partly on their importance in the co-ordinated plan of research, partly on the interest in or flair for any particular line of work which a member of the staff may develop, and partly of course on the time when sufficient data have been collected.

Although most of the scientific staff have spent considerably more time at home than at sea, it would not be possible for them to deal with every aspect of the work; but we have been greatly assisted by specialists in various institutions here and abroad, or working independently, to whom we have been able to farm out much of the material. Of these about forty have so far contributed one or more monographs to the *Discovery Reports*, and a number of others are examining various parts of the general collections. Many of these contributions are systematic reports (prepared in the style of the *Challenger Reports*) on various groups of marine animals; and although the material on which they are based was usually obtained with little interruption of the field programmes, they have added very considerably to what is known of the marine fauna. Other contributions from specialists include papers on such subjects as bottom sediments, special studies of invertebrates, the geology of

the Falkland Islands Dependencies, and so forth. All papers are published in the order in which they are received, and are distributed to many libraries at home and abroad, the balance being kept for sale and incidental presentations.

The work of the Discovery Committee is to be judged mainly by its contribution to oceanography and to the biology of whales, but the Committee has also had advisory functions. They have always been represented at international conferences on whaling, and have been able to give advice and assistance, not easily supplied from other sources, to other expeditions, departments and institutions.

The present and future

To review in a few words the ground which has been covered up to the present by the Discovery Investigations, I should say that four main points might be made. First, a survey—not quite complete yet—has been made of a large ocean area and the seasonal changes which take place in it. Secondly, I think a real step forward has been made in relating the oceanic fauna to the currents and water masses. Thirdly, researches on the biology of whales have helped the international regulation of whaling. And fourthly, a pool of information on the Southern Ocean and Antarctic regions has been built up.

I have emphasized the intermediate position of the Discovery Committee's work between an expedition and an institution, but it might be more correct to say that it has been in a state of transition, for it is now approaching the status of a permanent service. Since 1939 until recently the ships were on charter to the Ministry of Transport, but the work at home has been continued with a managing committee under the chairmanship of Mr J. M. Wordie, one of the original members of the Discovery Committee. In the National Institute of Oceanography, to which Dr Deacon has been appointed as director, we have now joined a larger organization which is to 'advance the science of oceanography in all its aspects', and as a going concern we can give it a flying start. This does not mean that the Discovery Investigations will lose their identity in the new Institute; in fact, the title will be retained at least for the continuation of the work in the south with which it is associated, and the *Discovery Reports* will continue to be issued. Under a provisional executive committee, acting for the Institute, our affairs, indeed, are moving fast. Additions have been made to the scientific staff, the *William Scoresby* is commissioned again, and the *Discovery II* will soon be ready for sea. During the first year or two the work of both ships will be in continuation of the Discovery Committee's programme, though some preliminary work in a wider field will be included. Thus the principal task of the *Discovery II* will be to round off the general survey of the Southern Ocean, but work will also be started in the central Indian Ocean. The marking of humpback whales during their winter migration off the north-west Australian coast will be the main item in the *William Scoresby's* programme, but she is starting with a preliminary survey of the Benguela current, and will test some possible methods of catching the famous Coelacanth fish, *Latimeria*, off East London. Although the ships are occupying most of our attention just now, we have in mind plenty of new developments in the more direct observations on whales.

Although much of the data to be collected in the near future must be properly comparable to what we have obtained in earlier years, we shall make use of some of the new instruments which have been developed more recently. The bathythermo-graph, for instance, can hardly be omitted now from the outfit of a research ship, and the *Discovery II* will also carry the Kullenberg piston core sampler and instruments for seismic sounding in bottom sediments, both of which offer many possibilities.

Deep-sea oceanography has travelled a long way since the time of the Challenger Expedition, and it seems inevitable now that a ship should specialize to some extent, at least within any given period or voyage. In our experience fruitful results are obtained from planned voyages in which the ship's principal movements are dictated by some major object, but which form a framework within which some attention can be paid to many more specialized problems. The major object might, of course, be a general survey of some ocean area, or it might concern a particular subject such as the exploration of the sea floor. The Discovery Investigations will in the future no doubt take their place along with other long-term researches in the physical and biological aspects of oceanography undertaken by the National Institute. We have been concerned over the years with a wide field of inquiry, ranging from problems in pure physical oceanography to such practical questions as the machinery of the regulation of whaling, and we have a flexible organization. It has all been a great adventure in the past, and I believe that it will be so in the future.

Wilkins Lecture

Robert Hooke

By E. N. da C. Andrade, F.R.S.

[Plates 12 to 14]

Science in England in the latter part of the seventeenth century is overshadowed by the mighty name of Newton, who has justly received the praises of all the great natural philosophers who came after him. In that springtime of science there were, however, in England a number of other men of genius who carried out work of prime importance—Robert Boyle; John Wallis and Isaac Barrow: Flamsteed and Halley; Willughby and Ray; Sydenham and Glisson; and Robert Hooke. Of these Robert Hooke has good claims to be considered the greatest. Probably the most inventive man who ever lived, and one of the ablest experimenters, he had a most acute mind and made astonishingly correct conjectures, based on reason, in all branches of physics. Physics, however, was far from being his only field: he is the founder of scientific meteorology; as an astronomer he has observations of great significance to his credit; he did fundamental work on combustion and respiration; he was one of the founders of modern geology. He has, moreover, a particular claim to the attention and respect of our Society, for from 1662 to 1677 he held the office of Curator and from 1677 to 1682 he was one of our Secretaries. He was always indefatigable in his services to the Society, and for a period he produced new experiments or discoveries at practically every meeting. Most writers who have really studied his work have given Hooke enthusiastic praise, yet, on account of certain difficulties of character—difficulties which he was not the only one to possess—his name does not seem to be honoured as it should be among men of science in general. No one has ever devoted a book to his life and achievements,* but he has been made the subject of casual and ill-considered criticism. It therefore seemed to me that it would be altogether fitting that I should attempt to recall to you something about this extraordinary man; about his services to science and his services to our Society.

Robert Hooke was born at Freshwater, in the Isle of Wight, on 18 July 1635, his father being curate of the parish.† Aubrey says that his father was one of the family of the Hookes of Hooke in Hants. Hooke was thus seven years older than

---

* The standard account is the *Life*, only some twenty-eight pages long, which Richard Waller prefixed to the *Posthumous Works* (1705). This is cited in further references as *Waller*, and the page given in Roman numerals, as used in the *Life*. The Life in Ward's *Lives of the Gresham Professors* (1740) is largely taken from this. Dr J. R. Morgan wrote a praiseworthy thesis on *The Contributions of Robert Hooke to the Physical Sciences*, in 1930, but this has never been published.

† No trace remains of the house of his birth, although it is traditionally believed to have been on the south side of a hill known to-day as Hook Hill.

Newton, a fact which probably had some influence on the relations between the two men. Like Newton he was a weakly child, but whereas Newton grew up strong and straight, Hooke was never physically sound. We have a description of him from Richard Waller, who was Secretary of our Society from 1687 to 1709 (as well as again at a later period) and must have known him well. He tells us that as to his person he was but despicable, which recalls Samuel Pepys' entry for 15 February 1664/5, where, after telling us that he was that day admitted to the Royal Society 'by signing a book and being taken by the hand of the President, my Lord Brouncker, and some words of admittance said to me', he adds, 'Above all, Mr Boyle was at the meeting, and above him Mr Hooke, who is the most, and promises the least, of any man in the world that ever I saw.' Hooke was very bent and crooked, but told Waller that he was straight until he was about 16, when he grew awry by working at the lathe. Hooke, who left some notes about his early life, also said that as a boy he was very sprightly and active in running and leaping 'tho' very weak as to any robust Exercise'. 'He went', says Waller, 'stooping and very fast having but a light Body to carry and a great deal of Spirits and Activity, especially in his Youth.' He was also, one gathers, meanly ugly, very pale and lean: 'His Eyes grey and full, with a sharp ingenious Look whilst younger; his Nose but thin, of moderate height and length; his Mouth meanly wide and upper Lip thin; his Chin sharp and Forehead large....He wore his own Hair of a dark Brown colour, very long and hanging neglected over his Face, uncut and lank.' Aubrey, who seems to have been his close friend and most anxious to speak well of him—he says that he was a person 'of great suavity and goodness'—also records that he was something crooked, that his head was large but the lower part of his face little and that his grey eyes were 'full and popping'. I think it important that you should know something of his appearance and great physical disabilities, and I quote so fully from the descriptions of those that knew him because I can say with some confidence that there is no known portrait of any kind of him, although in his diary* he seems to suggest that one Bonus (usually spelt Bownest), a known artist, drew his picture. It is one of my ambitions to find that picture.

The infirm boy, like Newton, showed a great taste for making mechanical models and for drawing. It is recorded that he was, as a boy, for some time an apprentice to Sir Peter Lely, the artist, but the smell of the oil colours provoked the headaches which worried him through his life. Whether he was actually apprenticed to Lely or no, he was certainly a competent draughtsman, as the drawing reproduced in plate 12 and his architectural drawings bear witness. We get a glimpse of him at Westminster, where he acquired a passable knowledge of Greek and Latin and studied mathematics. At Westminster he lived with the famous headmaster Dr Busby, and it is worth noting that not only did Busby, as far as one can judge, like and encourage him but that later in life he and Busby were on intimate terms and often dined together. Hooke had some firm friends.

In 1653, at the age of eighteen, he went to Oxford, with no private means. He had a chorister's place at Christ Church, 'which', says Aubrey, 'was pretty good maintenance'. The position of chorister was, however, a lowly one; he was also servitor

---

* Diary of 1672 to 1680, edited by the care of our Mr H. W. Robinson and Mr W. Adams.

to a certain Mr Goodman. This subordinate social position of Hooke's is to be remembered when we come to consider his status in the Royal Society.* It is generally held that Dr Thomas Willis, of Christ Church, who had equipped a laboratory, soon employed Hooke as assistant and recommended him to Robert Boyle. In any case it is as Boyle's assistant that Hooke's scientific career begins.

Oxford was at that time the great centre of the new experimental philosophy. From about 1648 onwards a brilliant band of men, including John Wallis, Seth Ward, William Petty, Thomas Willis and John Wilkins, used to meet to discuss the new method of experiment; after Petty and Wilkins had left Oxford, the meeting place was at Robert Boyle's house. Just when Hooke went to help Boyle is not known; if the young man whom John Wilkins recommended to Boyle in a letter was actually Hooke, the year was 1653.† It was about 1655, however, that he began to mix freely with the great men who were enthusiastically experimenting and that he began to be known for his genius in mechanical devices.

In 1658 or 1659 he made for Robert Boyle the first air pump used by Boyle in his experiment, which was also the first air pump made in England. It was not, of course, the first air pump, for in 1657 Schott had published an account of Guericke's pump, but it was a great improvement on this.‡ Hooke says: 'I contriv'd and perfected the Air Pump for Mr Boyle, having first seen a Contrivance for that purpose made for the same honorable Person by Mr Gratorix, which was too gross to perform any great matter.'§ Boyle, in his *New Experiments Physico-Mechanicall* (1660), acknowledges that Hooke 'fitted him with' a pump, after 'Mr. G.' had failed. This pump, although crude in design compared to the mechanical contrivances of Hooke's maturity, acted so well that the workmanship must have been excellent. About the time of the making of the pump, that is, 1658 or 1659, Hooke was much occupied by projects for flying. It is typical of Hooke's perspicacity that he abandoned these projects because he decided, as a result of trials and calculations, 'that the Muscles of a Man's Body were not sufficient to do anything considerable of that kind'.

From now on we are to be confronted with the difficulty of coping with the stream of inventions, notions, brilliant suggestions, accurate observations, daring speculations and prophetic conjectures that poured from Hooke's fertile brain and contriving hands. It will be impossible even to mention them all; to classify them will be difficult; in many cases, in view of the scanty record, it will be hard to decide exactly what was done. Practically everything, however, will bear witness to a truly extraordinary inventiveness and a truly modern outlook. Sometimes Hooke is wrong, but he is wrong in a strictly scientific and not a medieval way. Very often the ideas which he tumbled out in such profusion were taken by others; sometimes his findings were reached quite independently by others, which Hooke

* In some ways, as, for instance, in his lowly position at Oxford; in his leaving without a degree, which was subsequently granted; in the doctor's degree conferred late in life; in his physical disabilities, Hooke makes us think of Samuel Johnson.

† See L. T. More, *Life of Boyle*, pp. 79, 80.

‡ Guericke's own account, describing a somewhat more elaborate pump, first appeared in 1672.

§ Waller, p. iii.

found hard to believe. At every stage we are witnessing the workings of a mind so active, so fertile in expedients, so interrupted at every hour, at every endeavour, by the inrush of new concepts, new projects, that it is hard to disentangle his doings. Newton said that he made his discoveries by keeping the subject constantly before him and waiting until the first dawnings opened little by little into the full light. This Hooke was quite unable to do; he totally lacked Newton's powers of concentration. His mind was restless, continually disturbed by fresh ideas, but they were nearly all good and many were of first importance.

In 1661, while he was still with Boyle, appeared a little book of fifty small octavo pages, *An Attempt for the Explication of the Phenomena Observable in an Experiment Published by the Honorable Robert Boyle*, which dealt with the rise of liquids in capillary tubes and with surface-tension phenomena in general, for Hooke recognized that the spherical form of drops, the rise of water at the edge of a vessel and the rise in fine tubes were all aspects of the same general behaviour. His explanations were wrong: he put everything down to the pressure of one fluid on the other; thus the spherical drop was due to pressure of the air. In particular, the rise in the capillary tube was due to the pressure of the air being less within the fine tube than in the room at large. In this connexion he carried out experiments on the difficulty of forcing air through fine tubes—on the viscosity of air, as we should say. If you smile at the notion that the pressure within a fine open tube may be less than outside, remember that Newton accepted it.* Hooke's little tract is full of acute suggestions; he says that the capillary rise explains the rise of oil in the wick of a lamp and possibly the ascent of sap in trees, and sets down many accurate observations on the behaviour of small floating bodies, due to surface tension (which, needless to say, he does not so name). Clearly it is not possible here to discuss this minor work of Hooke's, but I cannot resist quoting the words in which he gives his method of work: 'For I neither conclude from one single Experiment, nor are the Experiments I make use of, all made upon one Subject: Nor wrest I any Experiment to make it *quadrare* [square] with any preconceiv'd Notion. But on the contrary, I endeavour to be conversant in all kind of Experiments, and all and every one of those Trials, I make the standards (as I may say) or Touchstones by which I try my former Notions. . . .' Thus the young man of twenty-six set down his scientific creed in an age when the experimental method was still making its way.

Now in these early days Hooke made an invention of capital importance, the history of which was most unfortunate for him, namely, the invention of the balance spring for watches. Until this time the portable spring-driven watch was governed by a swinging bar, which rebounded to and fro as the pallets struck the teeth of the escape wheel. This is clearly a very rough control; the error in the best of these watches was about a quarter of an hour a day.

* In 1675, in *An Hypothesis explaining the Properties of Light, discoursed of in my several Papers*, Newton says:-'as the air can pervade the bores of small glass pipes, but yet not so easily as if they were wider; and therefore stands at a greater degree of rarity than in the free aerial spaces, and at so much a greater degree of rarity as the pipe is smaller, as is known by the rising of water in such pipes to a much greater hight than the surface of the stagnating water, into which they are dipped.'

Hooke, whose interest in astronomy was awakened by Seth Ward, had been experimenting with pendulum clocks as accurate time-keepers and, according to his account, conceived the idea of springs as a control for watches, with the longitude problem in his mind. He asked for Boyle's help to protect his invention, Boyle invoked Lord Brouncker and Sir Robert Moray, and an agreement was drawn up somewhere about 1660 between Boyle, Brouncker, Moray and Hooke for the marketing and protecting of what was, in Hooke's words, 'a pocket watch, accomodated with a Spring, apply'd to the arbor of the Balance to regulate the motions thereof'. According to Hooke's subsequent account* the negotiations were broken off because he would not agree to a clause in the document drawn up by Moray stating that if, after Hooke had disclosed his invention, any other persons should find a way of improving the principles, they should have the benefit thereof during the time of the patent. Hooke contended that it was easy to vary his principles and perhaps to effect improvements (it being, as he said, *facile inventis addere*) and that this clause might deprive him of the benefit of his invention. The existence of this document is well attested. I do not think that an agreement of this kind would have been drawn up unless a going watch had been produced. Hooke was a young man without social position, protection or influence; the other three were not business men ignorant of science and anxious to make a lucky speculation, but critical spirits apprised of the science of their time. Robert Boyle is known to all as the author of the *Sceptical Chymist*; Brouncker was a first-rate mathematician and first President of the Royal Society; and Moray, if not of this calibre, had extensive scientific knowledge† and was a man of the world. Incidentally he held office as President unofficially at the time when the Society was in its primitive stage.

If I insist upon the significance of this document it is because the subsequent history of the spring-governed balance wheel is involved. Hooke says that he showed the watch described in the words already quoted to the three men interested in the patent, but there is no other record. In *The History of the Royal Society* by Thomas Sprat, dated 1667, is a list of instruments invented by the Fellows, given without the names of the originators, which contains several known inventions of Hooke's; in this is recorded 'Several new kinds of *Pendulum Watches* for the Pocket, wherein the motion is regulated by Springs or Weights, or Loadstones, or Flies moving very exactly regular.' The use of loadstones and 'watch flies' is known for certain to have been put forward by Hooke. A letter dated 30 September 1665, from Sir Robert Moray to Oldenburg, very definitely infers that Hooke had already applied a spring to the arbor of the balance.‡ I consider it practically certain that Hooke had invented the modern spring-controlled balance wheel by 1665, and that earlier he had a watch controlled by a straight spring attached to the balance wheel. I think it very probable that working watches had been made under his supervision —there is a tradition that he presented one to Wilkins, after whom this lecture is

* In *A Description of Helioscopes and some other Instruments*, 1676.

† 'Authorities concur in assigning to him an extensive knowledge of natural philosophy and mathematics', Weld, *History of the Royal Society*.

‡ This letter, from which Waller quotes on p. vi of the *Life*, is still in the possession of the Royal Society (M.1.13).

named, about 1661. He frequently affirmed that he had made a watch more exact than any pendulum clock, but does not seem to have produced it when requested by the Society to do so. The whole matter was brought to a head when Huygens, in 1674, brought out a watch with a balance wheel governed by a spiral spring. His application for a French patent failed because Hautefeuille claimed that he had already used springs in this way. Huygens's invention revived Hooke's interest in the balance wheel, he busied himself with Tompion making watches, and in 1675 he presented to Charles II a spring-controlled watch, inscribed *Rob. Hooke invenit 1658 T. Tompion fecit 1675*, which, of course, is no proof that he invented it in 1658. He referred his claim to priority to the Society, which did not support him, but seemed to favour Huygens. In this connexion it must be remembered that Oldenburg, the Secretary from 1663 to 1677, was an enemy of Hooke's, who undoubtedly made mischief between Hooke and Newton, and that further he had a financial interest in Hooke's discomfiture, for Huygens had assigned to Oldenburg his English patent rights, and Oldenburg tried in vain to get an English patent granted. But even Oldenburg, writing in 1675, admits ' 'Tis certain then, that the Describer of the Helioscope' (Hooke) 'some years ago caus'd to be actually made some Watches of this kind', adding, 'but it is as certain that none of these Watches succeeded', which he cannot have known. Hooke gave a circumstantial account of his invention in *A Description of Helioscopes*, published in 1676. The episode is of importance, since Hooke felt his lack of support from the Royal Society very deeply, and especially the part which Oldenburg played in it.

If I have devoted some time to this question of the balance wheel it is not so much because it is of capital importance—for Hooke made many discoveries of capital importance—but because it is an example of the difficulty of settling precisely the history of Hooke's inventions. He would come upon some brilliant and original mechanical conception, take it a certain distance, generally far enough to show that it would work, and then, stopping short of the final stage and precise publication that would convince the world of his priority, take up some other brilliant notion that had come into his head. He would then, naturally, feel aggrieved when someone else produced an elaboration or great extension of his work, generally without any acknowledgement. It is clearly impossible to discuss, however briefly, the history of each of his inventions. What I now propose is to describe briefly to you the course of his life and then to try to estimate his scientific achievement.

I have already referred to Hooke's life at Oxford as paid assistant to Robert Boyle. The Royal Society, having received its Charter in 1662, set about the appointment of a Curator whose duty was 'to furnish the Society every day they mett, with three or four considerable experiments, expecting no recompense till the Society gett a stock enabling them to give it'. Robert Hooke was appointed to the post. From this time onwards he devoted much of his time and much of his invention to the Society's interest; it is scarcely an exaggeration to say that without his experimental demonstrations, which over long periods took place weekly, the Royal Society would have died, as did the Accademia del Cimento, or declined into scientific insignificance, as did the Académie des Sciences on the death of Colbert in 1683.

To turn the pages of Birch's *History of the Royal Society*, which is a record, sometimes a very detailed record, of the meetings, week by week, from 1660 to 1687, is to convince oneself that Hooke's experiments and theoretical ideas, and the discussions which they provoked, were the main agent that held the Society together.

In 1664 it was voted that he should be given a regular salary, with the help of Sir John Cutler's promise of £50 a year. The record of Hooke's payments is, like everything else to do with him, scattered and complicated; in general it may be said that the payments were poor and irregular. Typical is that Cutler never paid what he promised, so that Hooke had to enter on a Chancery suit against Cutler's heirs to obtain his dues. In 1696, about thirty years after Cutler's first promise, the suit was determined in Hooke's favour, an event that gave him the greatest satisfaction. On one occasion, the Society proposed to pay Hooke in copies of Willughby's *Historia Piscium*, a work printed at the expense of the Society of which a large number of copies were left unsold.* Hooke apparently lived in the early days of the Society by consulting work, as we should call it now, done for instrument makers and technologists of all kinds. After the Great Fire he had lucrative employment in connexion with the rebuilding schemes.

In 1665 appeared his great work *Micrographia*, inscribed to Charles II in a short preface that displays Hooke as a writer of polished and beautiful prose. This book, as the title would suggest, contains a record of a large number of microscopical observations, made with the assistance of a compound microscope, which, judging by the plates of objects seen through it, must have been an excellent instrument. An original Hooke microscope, of the type described in the *Micrographia*, is shown in plate 13.

The plates in the *Micrographia* are beautiful in themselves, but also record a number of fundamental discoveries, to some of which I shall refer later. Sachs, the historian of botany, puts Hooke with Malpighi, Grew and Leeuwenhoek as 'endeavouring by earnest reflection to apply the powers of the mind to the objects seen with the assisted eye, to clear up the true nature of the microscopic objects, and to explain the secrets of their constitution'. The figures of the gnat, the flea and the louse were long famous. But microscopic pictures and their discussion form but a small part of the book. In it we find important theoretical discussions of the nature of light and heat, on which I shall animadvert later; further discussion of capillarity on the lines of his earlier tract; experiments on the thermal expansion of solids and liquids; shrewd speculations on tempering of metals; observations on crystal structure; astronomical discussions, including attempts to form artificially craters like those of the moon; and accounts of the magnitudes of stars, in which occurs the statement that more powerful telescopes would discover fresh stars: 'I am apt to think, that with longer Glasses, or such as would bear a bigger aperture, there might be discovered multitudes of other small Stars, yet inconspicuous. And, indeed, for the discovery of small Stars, the bigger the *aperture* be, the better

---

* It was the printing of this book that so exhausted the Society's funds that there was nothing available to pay for the printing of Newton's *Principia*; in consequence, Halley undertook to print it at his own charge. For payments in 'books of fishes' see Weld, *History of the Royal Society*, I, 310.

adapted is the Glass.' Further, we must note that the book contains a very full discussion of the colours of thin plates, such as flakes of mica, air films between glasses, and bubbles not only of soapy water, but of rosin and several other substances. These observations were the cause of subsequent dispute with Newton.

The *Micrographia* gained Hooke considerable fame at home and abroad. Pepys was delighted with it; it received a lengthy review in the *Journal des Sçavants*, with folding copies of two of the plates. Sturm in his *Collegium Experimentale*, published in 1676, praises it in the highest terms, saying that he has seen an English microscope and that whoever does not admit that England bears the palm for this kind of instrument is either envious or has not seen Hooke's *Micrographia*. He gives a folding plate of a Hooke microscope.

In 1665 the Great Plague caused the discontinuance of the weekly meetings of the Royal Society, and it appears that Hooke went with other Fellows to Banstead Downs in Surrey, where many experiments were made. After the Great Fire in 1666 Hooke's abilities as an architect and surveyor obtained for him the lucrative post of City Surveyor, which brought him into close contact with Sir Christopher Wren—in fact, it is not too much to say that Hooke was Wren's chief assistant in all matters concerning the rebuilding.* In Hooke's *Diary* for 1672 to 1680, Wren is mentioned on nearly every page. Very frequent, too, is the entry of a 'view' for 10 shillings, which refers to a survey of a house foundation. It is meet that I should here refer to Hooke's eminence as an architect. Among other buildings designed and built by him were Bedlam Hospital, the Haberdashers' Alms-Houses, and Montague House, praised by Evelyn: 'On the whole it is a fine palace, built after the French pavilion-way by Mr Hooke.' It stood on the site of the British Museum, and was burnt down in 1686. Incidentally, Hooke produced a method of making bricks, with less clay and more speed than had been done before, at the Royal Society in 1667, but I can find no more of it. He even prepared a model for rebuilding the city, which appears to have represented a plan by which all the chief streets should be laid out on a rectangular or 'gridiron' pattern, as in Mannheim and New York. It is characteristic that, while the Great Fire ceased on 12 September 1666, this model was laid before the Royal Society on 19 September. The Society was well pleased with it.

It was in 1672 that the unfortunate disputes with Newton took place concerning the experiments on the decomposition of white light. At the time of the preparation of the *Principia* they flared up again. I touched upon them when I had the honour of lecturing before you on Newton.† I do not, I think, underrate the difficulties that might easily occur in intercourse between two such men as Hooke, touchy and anxious for his priority, and Newton, secretive, self-centred, assured and desperately averse from all controversy. Nevertheless, each man found much to respect in the

---

* '...disposed him to take to his Assistance Mr. *Robert Hooke*, Professor of *Geometry* in *Gresham College*, to whom he assigned chiefly the Business of measuring, adjusting and setting out the Ground of the private Street-houses to the several Proprietors....' *Parentalia*, p. 263. There is evidence for the opinion that Hooke was responsible for the design of the Monument.

† *Newton Tercentenary Celebrations* of the Royal Society, 1942 and 1946.

other—Newton in certain letters refers to Hooke in a very complimentary manner and was, from his copious notes which still exist, a close student of the *Micrographia*. Hooke could not but admire Newton's prodigious experimental skill and knew, I am sure, well enough that he could never rival Newton as a mathematician or in quantitative theory. Hooke, furthermore, did not long bear rancour if left unprovoked; I have already pointed out how aggrieved he felt over Huygens's watch, yet later on he made very complimentary references to Huygens's work.

The trouble between the two men began with the publication, in February 1671/2, of Newton's first paper on the decomposition of white light by the prism. Newton was, at the time, unknown in London except for the reflecting telescope which he sent to the Royal Society in 1671. Hooke, the senior man, already firmly established as the author of the *Micrographia*, was concerned to defend his theory of colour and criticized Newton's interpretation of his experiments, although he admired the experiments. I have no doubt that Oldenburg, who hated Hooke—in my opinion as the man who dominated the Royal Society meetings where he wished to shine as the administrator on whom all depended—did all he could to make trouble between the two men. Hooke wrote a very civil letter to Newton in which he spoke of 'two hard-to-yield contenders put together by the ears by other's hands and incentives' and Newton wrote to Oldenburg 'Pray present my service to Mr Hooke, for I suppose there is nothing but misapprehension in what lately happened', but it is unlikely that Oldenburg ever conveyed the compliment; he was concerned to keep up the dispute. If only John Locke had been the intermediary!

All the time that he was working with Wren on the rebuilding of the city Hooke was actually engaged in experiment, although not quite such a regular attender at the Society meetings as before, which can be understood. The middle seventies saw him, however, particularly active scientifically. In 1674 he published *An Attempt to Prove the Motion of the Earth* with a very courteous dedication to Sir John Cutler, who had promised him emoluments which he never paid, and his *Animadversions on the First Part of the Machina Coelestis*, both works which I shall have occasion to consider later. In 1676 we have his *Description of Helioscopes*, with the motto *Sic vos non vobis*. *Lampas* and *Cometa* followed in successive years and in the same year as *Cometa*, 1678, appeared *Lectures de Potentia Restitutiva or of Spring*, which contains Hooke's law, the sole achievement of this fertile genius known to most young men of science. This tract contains a record of several other prime discoveries. All these were bound up together in 1679 as a book with the title *Lectiones Cutlerianae*, a further compliment to the bilking Cutler, who thus achieved immortality without cost to himself.

In 1677 Oldenburg died and Hooke was appointed to the Secretaryship of the Royal Society, which he gave up in 1682. In 1679, after the *Philosophical Transactions* had ceased to appear, Hooke, at the request of Council, undertook to publish a kind of scientific sheet mainly concerned with what was being done abroad. This publication, known as the *Philosophical Collections*, ran from 1679 to 1682. Hooke then handed over the Secretaryship to Robert Plot and the *Philosophical Transactions* again appeared. There is no reference to the *Philosophical Collections* in our *Record*—Hooke's usual fate.

It was as Secretary that, in 1679, Hooke approached Newton to ask for a scientific contribution to the Society in a letter which was not only courteous but showed a complete lack of any ill feeling. Newton replied in a letter that I have quoted before, saying that he had lost interest in science.* In this letter Newton made a slip about a falling body on a moving earth, and Hooke corrected him, which irritated Newton. In a letter of 6 January 1680, Hooke definitely suggested the inverse square law to Newton, who, it is at the present day plain, was already in possession of it, but Hooke could not know that. Hooke was likewise perfectly clear on the basic principles on which the problem of planetary motion was to be solved, as I shall show when I come to discuss Hooke's scientific performances, but lacked the mathematical power to prove Kepler's laws. To Hooke's very courteous letter of 17 January 1680, which once more put forward the inverse square law, and asked for Newton's 'thoughts', Newton returned no answer. It is understandable, then, that Hooke, who could not have known that Newton already had the outline of the planetary theory in his head when he was at Woolsthorpe as a young man, was irritated when he learnt, six years after this letter, that Newton was going to publish the solution of the problem of planetary motion, on the lines which he had put forward, without any acknowledgement to him. All that he wanted, according to Halley's letter to Newton, was some mention in the preface, and, in support of his claim to this, at the least it seems quite possible that Hooke's correspondence set Newton's mind on the matter again. A civil word would have cost Newton nothing. He replied to Halley in two letters bitterly attacking Hooke and made no mention of him either in the *Principia* or in the *Opticks* (the publication of which he delayed until after Hooke's death), although it is quite certain from his written notes that he had carefully studied the *Micrographia*. Many of us will wish, with More, the biographer of Newton '... that he (Newton), in the full plenitude of his fame, could have shown more tolerance and a greater sympathy for that brave mind and spirit, housed in a suffering body'.

After giving up the Secretaryship Hooke continued to carry out experiments and to address the Society on all possible subjects, including the nature of memory and the notion of time, as well as the nature of comets. Everything that remains of his is full of interest and astonishingly modern in spirit. From about 1681 onwards it appears that he seldom left any full account of his lectures which could be entered in the Society's register, but intended later to write them up himself and publish, which he never did. That we have, in his *Posthumous Works*, a printed account of so much that he did which was never published during his life is due to Richard Waller, who brought out the volume with great care, preferring to print the manuscript just as the author left it rather than to amend or abbreviate. The book is dedicated to Newton, as President of the Royal Society! A further small volume containing papers by Hooke was brought out by Derham in 1726.

In 1687 he was much distressed by the death of his niece, Grace Hooke, and apparently became from then on 'more unactive, melancholy and cynical'. Nevertheless, he continued to contrive; for instance, in 1691 he was busy with instruments for sounding the sea and he built Aske's Hospital (Haberdashers' Alms-House).

* *Newton Tercentenary Celebrations, 1942*, p. 10.

He continued to lecture intermittently, when he was well enough, until 1696, when he seems to have become more or less a permanent invalid. The next year his case appeared hopeless, and both he and his friends thought that he would die forthwith. However, he continued in a miserable state—I will spare you such details as are recorded—until, in March 1703, he died in a condition of complete exhaustion. At his funeral all the Fellows then in town attended, 'paying', says the recorder, 'the Respect due to his extraordinary Merit'. He was, it is set down, decently and handsomely interred in St Helen's Church, Bishopsgate, where Sir Thomas Gresham also lies buried. The site of his grave is unmarked and unrecorded.
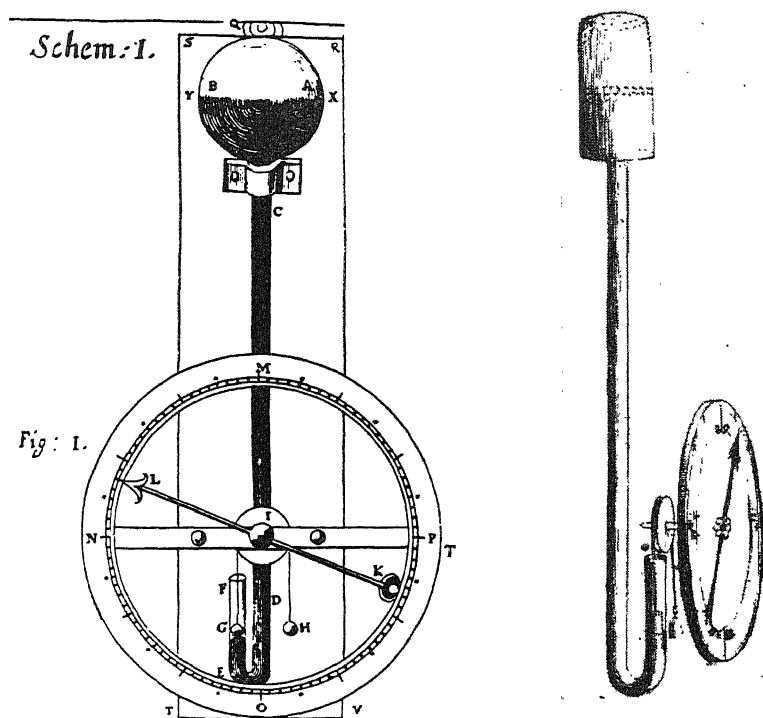

FIGURE 1. Hooke's wheel barometer: on the left is the diagram as given in the *Micrographia*; on the right Hooke's original sketch, in the possession of the Royal Society.

Let us now turn to Hooke's scientific achievement and consider first what Hooke did for the scientific instrument. It is in this field that he undoubtedly excels any other figure in the history of science. First of all, he invented all the meteorological instruments. His wheel barometer, shown in figure 1, is familiar to everybody; such barometers are still in use. Practically the last thing on which he busied himself was his marine barometer. As regards the thermometer, he proposed and used the freezing-point of water, 'common distilled water, that is so cold that it just begins to freeze and shoot into flakes', as the standard zero, the lower fixed point, just as it is used to-day. He had no upper fixed point, but avoided the necessity for one by a standard method of calibration. He prepared a cylindrical vessel, with a tube of one-tenth its diameter protruding from it, and marked on this tube graduations

representing one-thousandth of the volume of the vessel (figure 2). Using the same liquid in this as in his thermometer, he could clearly calibrate his thermometer in divisions representing an expansion of one-thousandth, and subdivide these. It was a properly standardized instrument that he made. His wind gauge was a plate pivoted so as to be blown aside by the wind, the strength of which was measured by the angular deflexion (figure 3). Instruments of this type were issued in the first
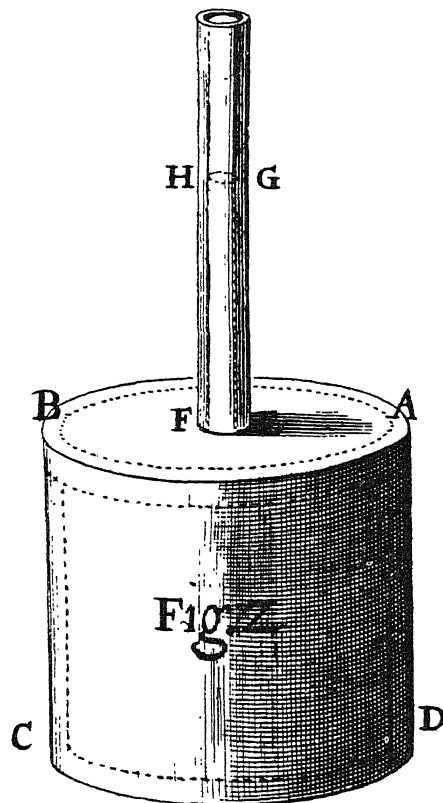

FIGURE 2. Hooke's vessel for use in calibrating thermometers. (From the *Micrographia*.)

world war. For measuring humidity, 'the dryness or moistness of the air', he employed the beard of a wild oat, which is naturally bent near the middle; the upper part rotates when the beard is moistened, owing to a natural twist in the structure of the lower part (figure 4). This beard he arranged so as to turn an index, which gave him his hygrometer. He also constructed a self-measuring rain gauge. He thus had at his disposal all the usual meteorological instruments, with the exception of a sunshine recorder. Further, following, as he freely acknowledged, the first design of Wren, he made an instrument which actually recorded the meteorological quantities. In this 'weather clock', which occupied him many years in the making, the instruments recorded every quarter of an hour on a paper strip by means of punches, the paper being carried on a rotating cylinder driven by a strong pendulum clock. According to the description, the barometer, the thermometer, the hygroscope, the rain bucket, and the wind vane all registered, the wind vane, of the
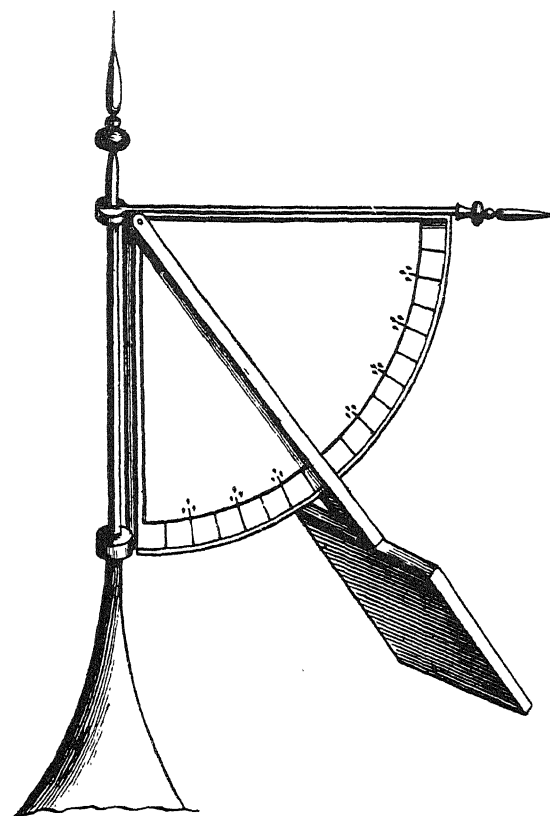
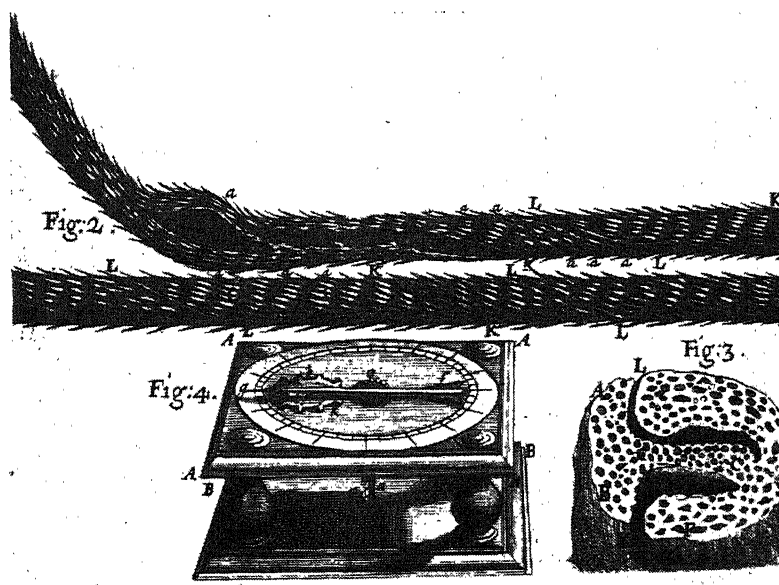FIGURE 3. Hooke's wind gauge.

FIGURE 4. Hooke's hygrometer. In 'Fig: 2' is shown the 'wreathed', or twisted, beard of the wild oat, in two parts: in 'Fig: 3' a cross-section of the beard; in 'Fig: 4' the hygrometer, with the beard mounted vertically so as to turn the index.

revolving type, recording the revolutions in 10,000's, 1000's and 100's. The direction of the wind was also recorded.* The Society saw the clock on 29 May 1679, but there is no detailed description.† On 2 April 1684 Hooke was told to help his assistant 'to reduce into writing some of the first papers marked by the weather-clock, that thereby the Society might have a specimen of the weather-clock's performances before they proceed to the repairing it'.‡ There thus seems no doubt that the clock worked, if but intermittently. But what a project even if it had not worked!

Hooke can safely claim to be the founder of modern meteorology. Typical of his quantitative spirit is his calculation, in 1662, of the height of the atmosphere from Boyle's law. He worked out that at 25 miles the pressure should be 0·02 atmosphere, which is, of course, much too high, but his conception of pressure decreasing exponentially is modern in spirit. The height of the atmosphere, he said, seemed to be indefinite, but might be many hundreds of miles. He made extensive barometric observations, with the object of finding the connexion between barometric pressure and weather; for instance, in 1677 he said that in sixteen years' observation he had found the quicksilver always to be very low, and to fall to that position rapidly, before any considerable storm, and therefore he hoped that the instrument would be of great use at sea, to foretell storms. He was perfectly clear as to the general polar circulation, saying that the air near the earth moved from the poles to the equinox and the higher parts of the air from the equinox to the poles. But perhaps what is most striking is the way in which he proposed systematic meteorological observations by means of his instruments. The table from his article on 'A Method of Making a History of the Weather', published in 1667 in Sprat's *History of the Royal Society*, is astonishing (figure 5). He is further quite clear that the heat of the sun is conveyed to the air by radiation falling on the earth. He thought of the weather as a matter of physical laws applied to the earth and its atmosphere, the movements of the air as governed by the absorption of solar heat and by the rotation of the earth, in quite a modern spirit. Here we may note that in 1678, pointing out that the rotation of a planet would lead to a spheroidal shape, he remarked on the influence of the diurnal motion of the earth 'which compounded with that of the moon he conceived to be the cause of the tides'.§ Proudman says of Newton's work on the tides: 'The only important factor which he did not mention is the dynamical effect of the earth's rotation.' Of course, there is no comparison between a considered mathematical theory and a shrewd qualitative generalization, but the remark does illustrate Hooke's exceptional acuteness.

Hooke constructed the first modern astronomical instruments. In his keen analysis of the accuracy attainable with astronomical instruments designed for measuring angles, Hooke at once met with the question of the resolving power of the human eye, which he was certainly the first to discuss. He goes into the matter in detail in his Cutlerian lecture *Animadversions on the First Part of the Machina Coelestis*, published in 1674. Hevelius, the author of the *Machina Coelestis*, had built

* *Philosophical Experiments and Observations of the late Eminent Dr Robert Hooke. Publish'd by W. Derham, F.R.S.*, 1726, p. 41.
† Birch, III, 487.
‡ Birch, IV, 277.
§ Birch. III. 390. See also Waller, xx.

an observatory at Danzig, and this he had equipped at great expense with magnificent open-sight instruments, which he used for all his measurements of angular distance. Those instruments were large and of beautiful workmanship: one is represented in figure 6. Hooke's criticism was that open-sights were inherently incapable of the accuracy that could easily be obtained with telescopic sights, at, he said, a tenth the cost. 'I am', he said, 'not satisfied that his Instruments are capable of making Observations more accurately than those of *Ticho*, though 'tis possible that they

A
S C H E M E
At one View reprefenting to the Eye the Obfervations of the Weather for a Month.

Dayes of the Month and place of the Sun. Remarkable houfe.	Age and fign of the Moon at Noon.	The Quarters of the Wind and its ftrength.	The Degrees of Heat and Cold.	The Degrees of Dryness and Moyfture.	The Degrees of Preffure.	The Faces or vifible appearances of the Sky.	The Notableft Effects.	General Deductions to be made after the fide is fitted with Obfervations: As,
14 II 12.46 / 4 8 12♉9.46. 4 8 Perigeu. 12	27	W. 2.|9 3|12 3½|16 10 W.SW.1|7	12 2 2	5|29 9 29 2 29		Clear blew, but yellowifh in the N. E. Clowded toward the S. Checker'd blew.	A great dew. Thunder, far to fhe South. A very great Tide.	From ' the laft quarter of theMoon to the change the weather was very temperate but cold for the fea-fon; the Wind pretty conftant between N. and W.
15 II 13.40 / 8 4 6 ♉24.51.N. 10	28	N.W. 3|9 4 28 1|7	2 8½|29 1/16 2 9 2 10 29			A clear Sky all day, but a little chec-ker'd at 4. P.M. at Sun-fet red and hazy.	Not by much fo big a Tide as yefterday. Thunder in the North.	A little before the laft great Wind, and till the Wind rofe at its higheft, the Quickfilver con-tinued defcend-ing till it came
.16 II 14.37 / 10 N.Moon. at 7. 25' A.M. ♊ 10. 8.		S. 1|10	1 10 28 ½			Overcaft and very lowr-ing.	No dew upon the ground, but very much upon Marble ftones, &c.	very low; after which it began to reafcend,
&c.	&c.	&c.	&c.	&c.	&c.	&c.	&c.	&c.

FIGURE 5. From Hooke's paper on 'A Method of Making a History of the Weather'.

may do it with somewhat less trouble and inconvenience', Ticho being, of course, Tycho Brahé. He gives half a minute as the limit of accuracy possible in measuring distances in the heavens with the naked eye 'and hardly one of a hundred can distinguish a minute'. The resolving power of the eye for two near bright objects is taken to-day as just under a minute (50 sec.), and Hooke demonstrated this before the Royal Society on 15 January 1673/4. The ability to set on an object may be slightly better than this—I do not know if it has been measured—but it certainly has a limit of about this magnitude. Hooke points out that it is useless to graduate an instrument with open-sights to take altitudes to seconds—quite apart, as he says, from the fact that reading to seconds does not mean accuracy to seconds—which Hevelius claimed to do by means of a vernier. He then discusses, in a quite modern

manner, the advantages of telescopic sights and describes in detail a mural quadrant with a fixed, horizontal telescope, for sighting on reference marks, and a movable telescope attached to the rotating arm. The telescopes (figure 7) are
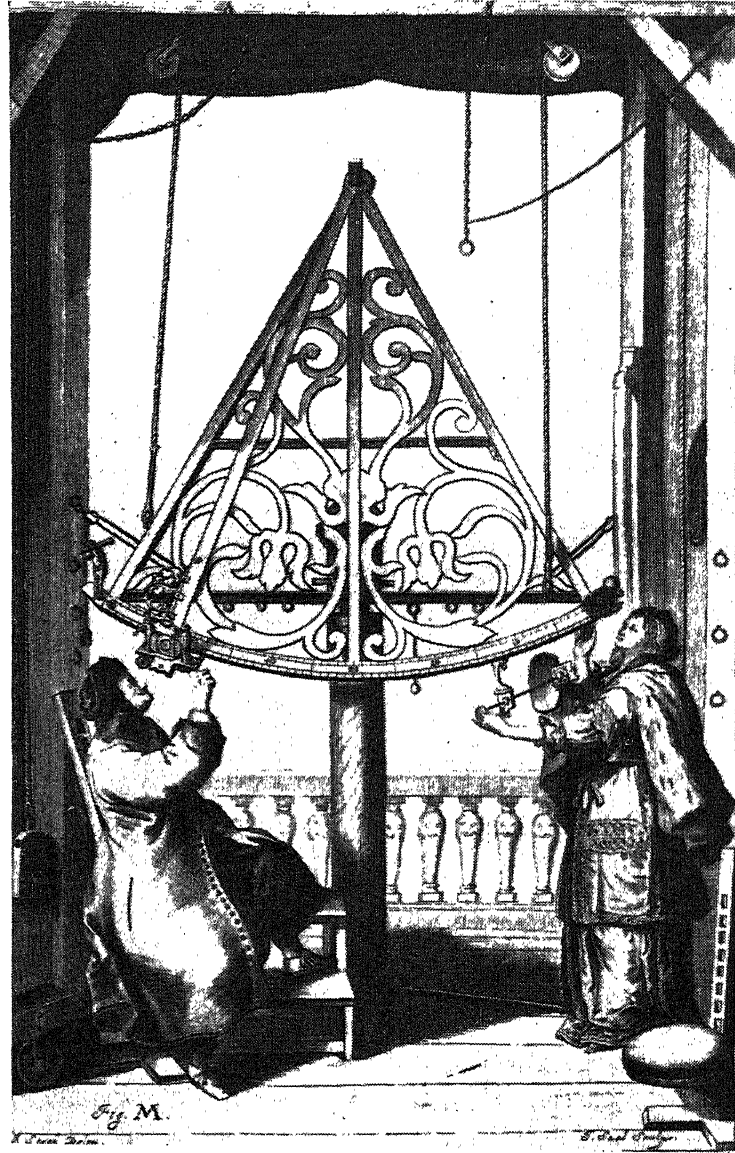
FIGURE 6. One of the open-sight instruments of Hevelius, contemporary with Hooke's telescopic sights.

furnished with cross-wires, made of spider's web or single silk fibre, at the focus of objective and eyepiece; furthermore, by metal reflectors, things are so arranged that an observer looking horizontally can see both objects on the cross-wires and bring them into coincidence by turning the moving telescope. Apart from telescopic sights, this is an anticipation of the Hadley (or Newtonian) sextant. As regards the

quadrant, Hooke describes how this may be accurately graduated by means of a screw—the first dividing engine (see figure 8). Incidentally, he describes the new invention of the universal joint or Hooke's joint (figure 9). He also describes a spirit
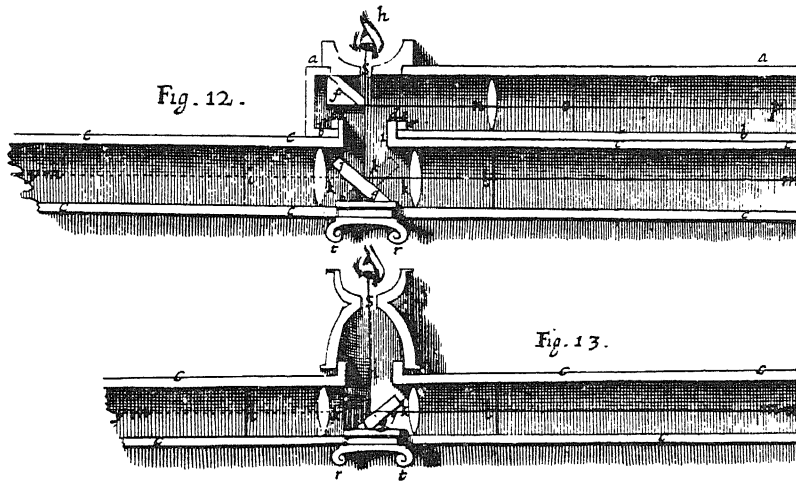
FIGURE 7. Hooke's telescopic sights, fixed and rotating, for his mural quadrant.
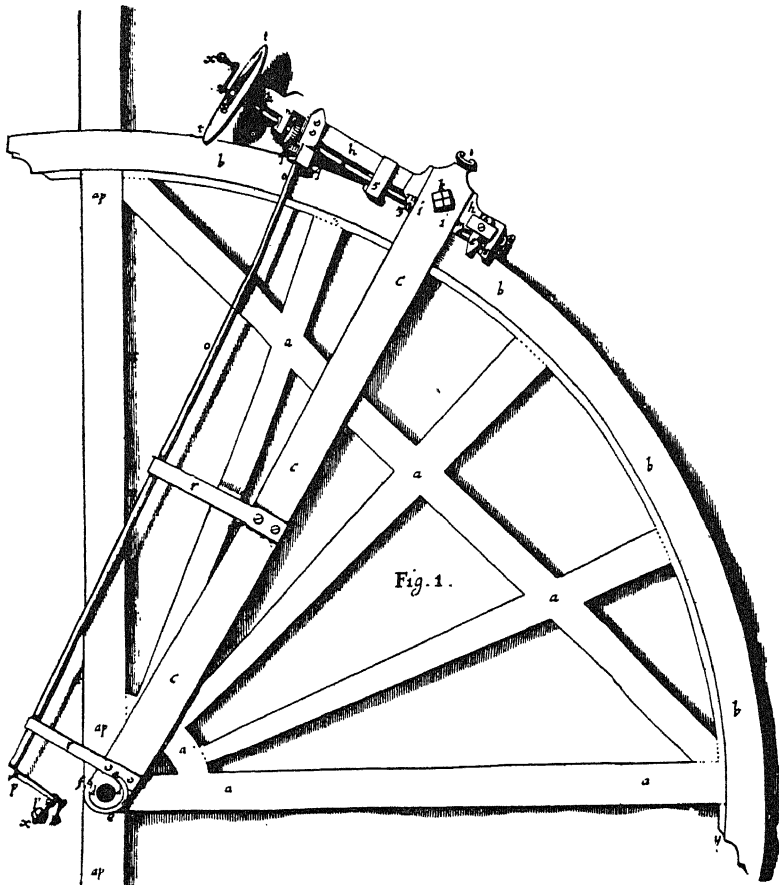
FIGURE 8. Hooke's quadrant, graduated by means of a screw.

level, but here he had been anticipated by Thevenot. Anyone who contrasts Hooke's instrument with those of Hevelius will see a difference of kind; one is completely modern in spirit, with a careful discussion of the limits of accuracy and of experimental error, the other is a much improved version of the instruments of the ancients.

But this is by no means all in this one section of the Cutlerian lectures. Hooke describes the first clock-driven telescope (figure 10), although there seems to be no evidence that he ever made such an instrument. It is driven by a conical pendulum
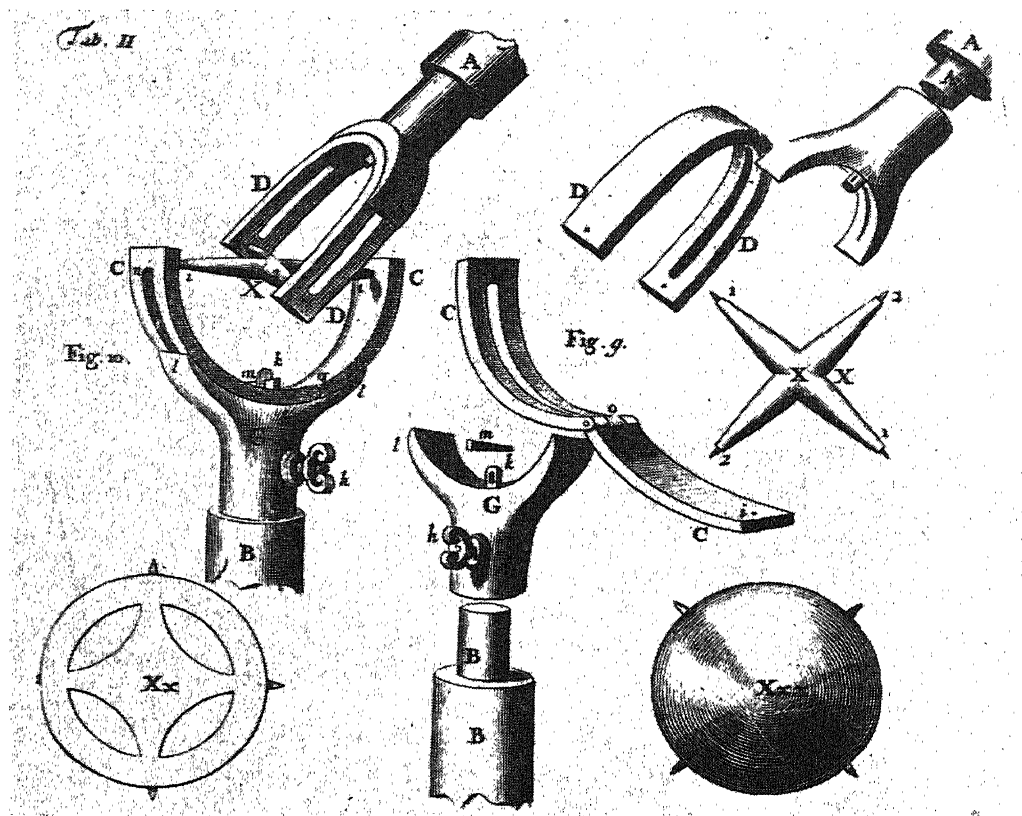
FIGURE 9. The universal joint or Hooke's joint.

(or circular pendulum, as Hooke called it), and the details given certainly read as if set down in the light of experience in the construction. As regards the conical pendulum, Hooke complains that Huygens, in his *Horologium Oscillatorium* (1673), gives a description of it without acknowledging that he, Hooke, invented it. This claim of his appears to be well-founded, for in Birch's *History* we find that on 23 May 1666 Hooke did discuss the conical pendulum. Casually, in the *Animadversions*, he gives the first invention of a spiral gear.

With regard to horology, I have already referred to the spring-controlled balance wheel, which I am satisfied to attribute to him, but about which there has been controversy. The invention of the anchor escapement is generally ascribed to Hooke, but here I can find no really satisfactory evidence for the attribution.

He constructed the first reflecting telescope of the Gregorian pattern, using spherical mirrors, after Newton had made one of the pattern to which he gave his name. It appears from Dr Pell's collection in the British Musem,* as well as from what Hooke himself says in his *Animadversions on Hevelius*, that Hooke invented about 1670 an engine for multiplying and dividing.† Samuel Morland published the account of his calculating machine in 1672. As we have seen, Hooke was responsible for the construction of Boyle's first pump.

Let us now look at a few of his minor inventions. He can certainly claim the iris diaphragm; the description in Birch, under the meeting of 27 July 1681, is precise: 'Mr Hooke showed his new-contrived aperture for long telescopes, which would
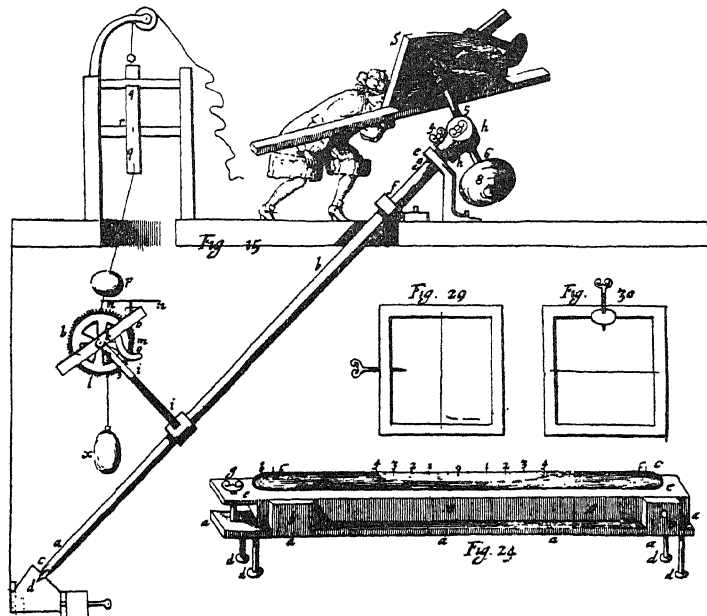
FIGURE 10. Hooke's figure of a clock-driven telescope.

open and close just like the pupil of a man's eye, leaving a round hole in the middle of the glass of any size desired; which was well approved of.' Apparently Hooke used to have the notion of some such device, make it in a morning, show it at the Royal Society, and forget it. At the same meeting he showed his improved helioscope, 'which would exhibit the figure of the sun very perfectly'. He invented ingenious machines for sounding the sea, and one for bringing up samples of sea water from the depths. He first used immersion lenses for the microscope, pointing out their advantage. He made the first refractometer for liquids.

His glimpses into the future are astonishing. I have already referred to his conclusions about flying. He says: 'The way of flying in the Air seems principally impracticable, by reason of the want of strength in humane muscles; if therefore

* See *A Brief Account of the Life, Writings and Inventions of Sir Samuel Morland*, by James Orchard Halliwell, 1838, p. 13.

† See also Waller, xix.

that could be supplied, it were, I think, easie to make twenty contrivancies to perform the office of Wings.' He therefore attempted to make artificial muscles, using a chain of small bladders, by blowing into which an immediate contraction could be produced. This does not appear to have been taken further. In 'a conjecture,
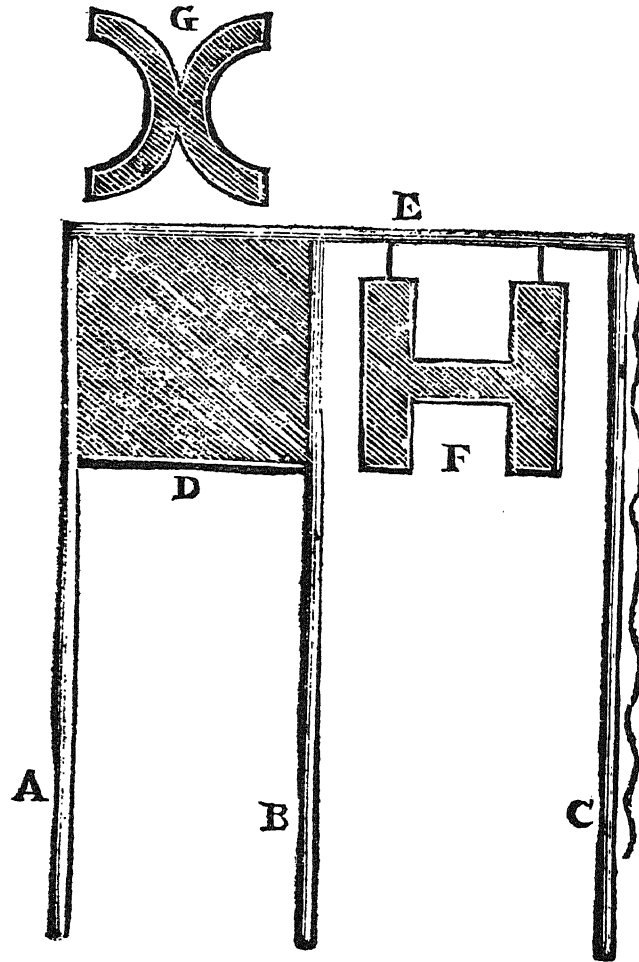
The Sentences, to be exprefs'd by one Chara&ter, may be fuch as thefe, in *Fig.* 2.

O *I am ready to communicate.*)(*I am ready to obferve.* (*I fhall be ready prefently.*

) *I fee plainly what you fhew.* _/ *Shew the laft again.* ⌒ *Not too faft. Shew fafter. Anfwer me prefently.* Dixi. *Make Hafte to communicate this to the next Correfpondent. I ftay for an Anfwer* ; and the like.

FIGURE 11. Hooke's figure of his optical telegraph; under it is shown a passage in which he describes his characters designed to represent sentences frequently needed.

that it may perhaps be possible to spin a kind of artificial Silk out of some glutinous substance that may equalize natural Silk', he says, 'And I have often thought, that probably there might be a way found out, to make an artificial glutinous composition, much resembling, if not full as good, nay better, then that Excrement, or whatever other substance it be out of which, the Silk-worm wire-draws his clew. If such a composition were found, it were certainly an easie matter to find very quick ways of drawing it out into small wires for use. I need not mention the use of such an Invention, nor the benefit that is likely to accrue to the finder, they being sufficiently obvious. This hint therefore, may, I hope, give some Ingenious inquisitive Person an occasion of making some trials, which if successfull, I have my aim, and I suppose he will have no occasion to be displeas'd.' It is possible that he anticipated the stethoscope. He says: 'There may also be a possibility of discovering the internal motions and actions of bodies...whether animal, vegetable, or mineral, by the sound they make; that one may discover the works performed in the several offices and shops of a man's body, and thereby discover what instrument or engine is out of order...I have been able to hear very plainly the beating of a man's heart.' He certainly invented the optical telegraph, subsequently developed, at the end of the eighteenth century, by Claude Chappe. He pointed out that the telescope enabled the signals to be read from a great distance, discussed conditions for good visibility and devised a system of signals for day and night use, all in a very practical spirit, as the passage reproduced in figure 11 will attest. These are a few examples of his sagacious anticipation.

If we turn from particular instruments and inventions to Hooke's general contributions to science, Hooke's law occurs at once to everybody. It was published in 1678 in *Lectures de Potentia Restitutiva or of Spring*, but had been announced two years earlier in the form of an anagram, *ceiiinossssttuu*; letters which, rearranged, make the words *ut tensio sic vis*—'as the extension, so the force', or, as we say nowadays, strain is proportional to stress. The law is illustrated in a number of ways—by the wire under axial tension, the helical spring, the spiral spring of the watch balance, the wooden cantilever (see figure 12). Hooke also quotes what is generally known as Boyle's law as a further example. Accurate measurements of the relation between volume and pressure of air he had published thirteen years earlier in the *Micrographia*; there he says that he first carried out experiments of this kind in 1660 and repeated them in 1661, when, having heard Townley's hypothesis, he modified the method of recording, and exactly verified that 'the Elater of the Air is reciprocal to its extension, or at least very neer', the elater being, in modern parlance, the elasticity. Thus it was Hooke who really discovered Boyle's law. That this is so is, I think, substantiated by two arguments: first, that Hooke, who always expressed the greatest veneration for Boyle, would never have published this in 1665 if it was likely to give pain to, or be disputed by, Boyle; and secondly, that Boyle says in 1662* 'wherein I had the assistance of the same person that I took notice of in the former Chapter...whom I had rather make mention of on this occasion, because when he first heard me speak of Mr Townley's suppositions about the proportion wherein Air loses its Spring by Dilatation, he told me he had the

---

* *A Defence of the Doctrine touching the Spring and Weight of the Air*, p. 64.

year before...made observations to the same purpose, which he acknowledged to agree well with Mr Townley's theory.' Thus Boyle's account and Hooke's are substantially the same. L. T. More, in his life of Boyle, says: 'The specific experiments point to Boyle, but the method of attack, which involved the physical and mechanical properties of the air, shows the influence of Hooke', and points out that after Hooke left him Boyle never did any quantitative work.
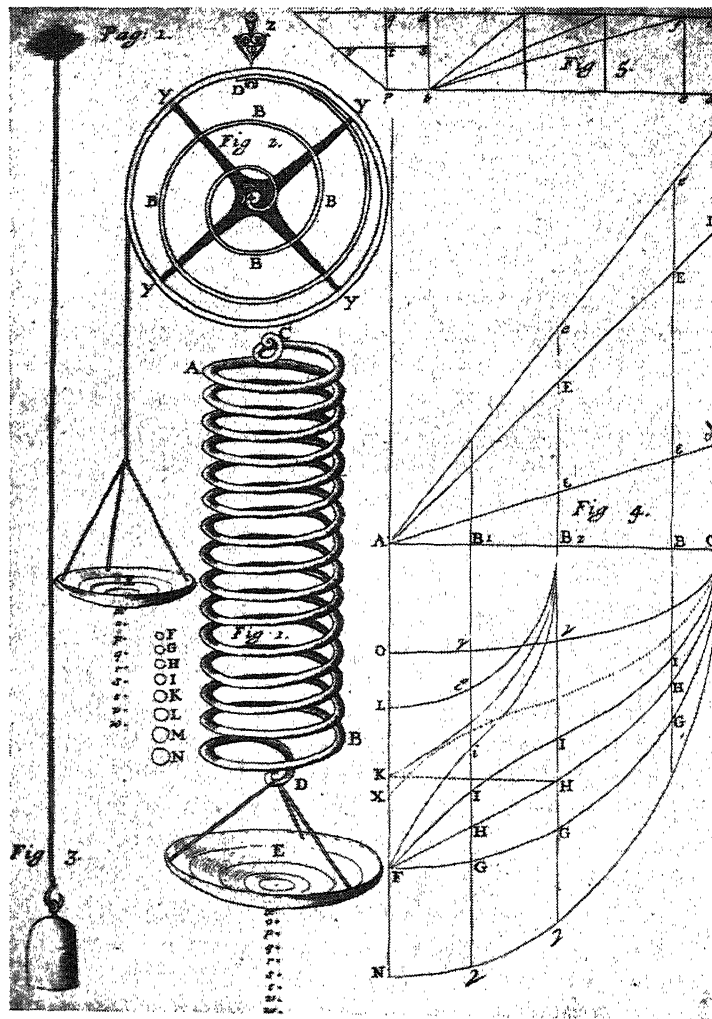
FIGURE 12. The plate which illustrates Hooke's *De Potentia Restitutiva or of Spring*.

The *Potentia Restitutiva* contains other things of major importance besides Hooke's law. Hooke shows that any vibration in which the restoring force is proportional to the displacement must be isochronous, a capital discovery. He points out that the restoring force of a bent beam is due to the extension of the under part and the compression of the upper part. Further, here and elsewhere, he develops a crude kinetic theory of matter. He supposes that all the particles of the universe are in

continual motion: 'Two or more of these particles joyned immediately together, and coalescing into one become of another nature, and receptive of another degree of motion and Vibration, and make a compounded particle differing in nature from each of the other particles'*—not a bad definition of molecules. He further supposes a universal ether which is the medium that 'conveys all Homogenious or Harmonical motions from body to body'. In a solid the particles touch at their extreme librations; in a liquid they do not, and the subtle ether gets between them. He says that all the particles of a solid are kept together by the motions of the surrounding ether and tend to fly apart by virtue of their own motions, and that 'according to the prevalence of the one or the other is the body more or less fluid or solid'. Air 'consists of the same particles single and separated, of which water and other fluids do, conjoyned and compounded', only they are very widely spaced—'its Vibrative Spaces exceeding large, comparative to the Vibrative Spaces of other terrestial bodies'.† He mentions a million a second as the kind of frequency he has in mind. On the basis of this general kinetic theory he explains the elasticity of all bodies. Jeans‡ says of this work: 'He suggested that the elasticity of gases resulted from the impact of hard independent particles on the substance which enclosed it, and even tried to explain Boyle's law on this basis.' When we turn to consider Hooke's thoughts on heat we shall again find this conception of universal vibration.

Hooke was the first to speculate in a modern way about crystal structure. In the *Micrographia* he points out that the forms in which alum and rock salt crystallize can be built up of spherical particles, and he gives examples, in two-dimensional drawings (figure 13), saying that he has built up models with 'globular bullets' of all the regular figures commonly met with in crystallization. He cites alum and rock salt—cubic crystals—in particular, but outlines a general programme concerning crystallization which he would like to pursue if he had time and assistance, ending: 'So that knowing what is the form of Inanimate or Mineral bodies, we shall be the better able to proceed in our next Enquiry after the forms of Vegetative bodies; and last of all, of Animate ones, that seeming to be the highest step of natural knowledge that the mind of man is capable of.' Our biophysical friends will agree.

Hooke was certainly one of the first, if not the first, to draw attention to thermal expansion as a general property of matter. 'This property of Expansion with Heat, and contraction with Cold, is not peculiar to Liquors only, but to all kinds of solid Bodies also, especially Metals.' As regards the nature of heat, it was, he said in the *Micrographia*, 'nothing else but a very *brisk* and *vehement agitation* of the parts of the body' and again '*Heat is a property of a body arising from the motion or agitation of its parts*'. He points out that if you rub any hard bodies together they grow hot, and says that a violent enough motion will melt a body, quoting, as an example, sparks struck from steel, which he showed microscopically to be particles molten

---

* *Potentia Restitutiva*, p. 9.

† *Potentia Restitutiva*, pp. 15, 16.

‡ Jeans, *An introduction to the kinetic theory of gases*, 1940, p. 3. See further Tait, 'Hooke's anticipation of the kinetic theory', *Proc. Roy. Soc. Edinb.* 16 March 1885. *Tait's Collected Works*, ii, 122.

int o a spherical form. Bodies that melt easily have parts unapt to cohere, so that a small degree of agitation keeps them fluid. Much else that he writes has a very modern sound, much more so than the speculations of the eighteenth century. The same we find when we turn to his work on combustion.

In the theory of combustion and respiration Hooke rendered services which have been often overlooked, although Thomas Thomson, in his *History of the Royal Society*, 1812, says: 'Mayow, in his Essays, published at Oxford about ten years after the *Micrographia*, embraced the hypothesis of Dr Hooke, without acknowledgement; but clogged it with so many absurd additions of his own, as greatly to obscure its lustre and diminish its beauty.' Ripe scholars like Dr McKie have of
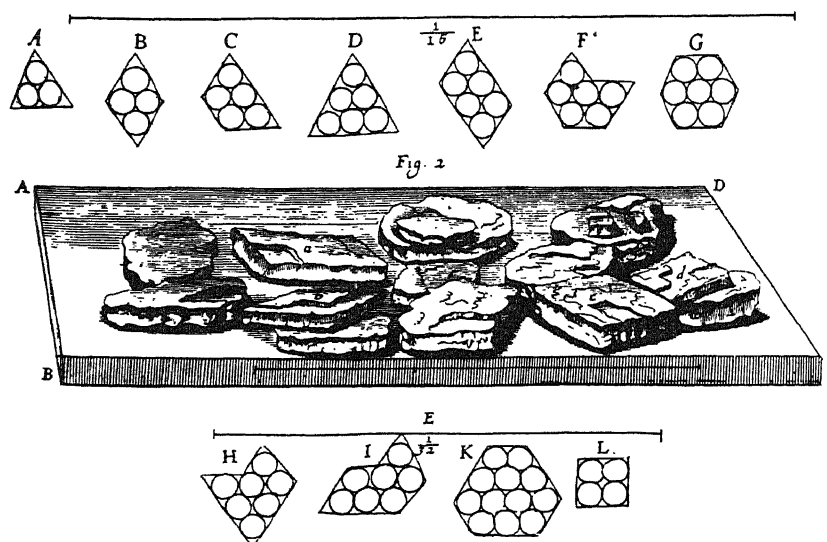

FIGURE 13. Hooke's representation, by means of spherical bullets,
of simple crystalline forms shown by alum.

recent years given balanced and appreciative accounts of Hooke's work in this field, but in the average history of chemistry Hooke's contributions are neglected. The record of the subject is confused and complicated, and it is possible that to abbreviate it is to introduce an artificial simplicity, but it must be attempted.

It was Boyle who first showed, with the pump which Hooke made for him, that reducing the pressure reduced the time that a candle could burn or a mouse or bird live in a given vessel. Hooke was responsible for the experiment, described in histories of medicine, in which a dog, of which the thorax and belly were displayed, was kept alive for a long time by air supplied through the windpipe by a pair of bellows, showing that the air, and not the muscular motion of the ribs, was the essential. In the *Micrographia* he described a crucial experiment in which he showed that wood heated in a closed iron vessel became charred, but would not burn in the absence of air. His theory of combustion was that combustible—sulphureous, as he called them—bodies dissolved in the air by virtue of a substance present in the air that was 'like, if not the very same, with that which is fixed in *Salt-peter*, which by multitudes of Experiments that may be made with *Salt-peter*, will, I think, most evidently be demon-

strated'. This substance was used up in combustion, so that a fresh supply of air, as, for instance, that produced by bellows, was necessary for continuing combustion. Thus he clearly realized the functions of oxygen without naming it further than by saying it was copiously contained in saltpetre, or, as he himself said very plainly, concerning respiration, air contains 'a kind of nitreous quality which being spent made the air unfit for respiration'. He further said that the function of respiration was to introduce from the air into the blood something essential to life and to remove from it something noisome that was discharged back into the air. Experiments to find out if the volume of air changed due to combustion were not decisive; it was here that Mayow, performing the experiment over water, made an advance. But I hope that I have indicated that, before his time, Hooke had arrived at the essentials of the true theory of combustion and respiration. Incidentally in 1671
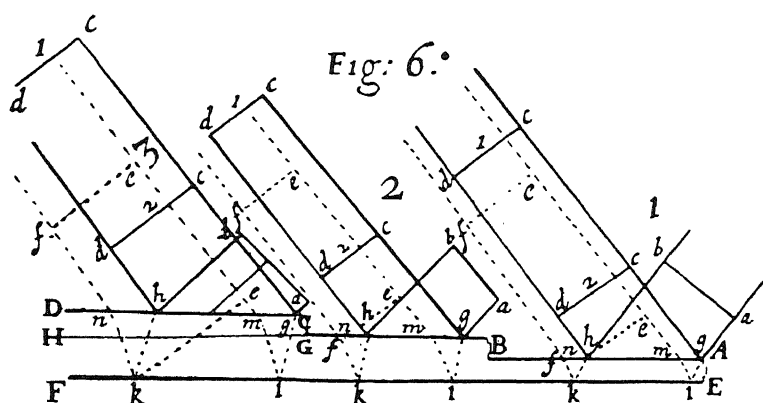

FIGURE 14. Hooke's diagram to represent the way in which the colours of thin plates are formed.

Hooke tried some experiments, with himself as the subject, on breathing at reduced pressure. The experimental compartment consisted of two tuns, one within the other, the space between being filled with water. He managed to retain the pressure at a quarter of an atmosphere below normal and stayed in this rarefied air for a quarter of an hour; now and then fresh air was let in. He felt little inconvenience except about the ears, as anyone who has flown at about 8000 feet can well believe. These must be the first experiments ever carried out on the effect of artificial changes of air pressure on human comfort.

As regards light, the *Micrographia* contains, first, a discussion of the colour of thin plates which shows Hooke's usual acuteness of observation and perception of essentials. Having found that the colour of flakes of mica only appeared when the thickness was very small, he produced colours with thin films of air between glass plates and observed the changes of colour that took place as the plates were pressed closer together. He was quite clear that the colours of thin plates were produced by the light reflected from the back of the plate acting together with that reflected from the front—his diagram (figure 14) would do for to-day. It is only his notion of the mechanism of interaction that is obscure. He treated the colours of thin bubbles

of various kinds. He observed with a lens pressed on glass the rings which bear Newton's name. He said that the colours of steel were produced by a thin lamina of 'vitrified' (oxidized) steel, and conjectured that the hardness of tempered steel was due to 'vitrified' substances interspersed through the pores of the steel. What I wish to stress here is not the possibility of interpreting this as a foreshadowing of modern theory but that Hooke is attempting what, I believe, nobody had considered before, the explanation of the change of properties of metals by heat treatment on strictly physical and structural lines. 'Those metals which are not so apt to vitrifie, do not acquire any hardness by quenching in water, as Silver, Gold, &c.' I should like to draw your attention to the fact that round about the beginning of 1676 Newton often refers to the *Micrographia*, and in his discourse read before the Society on 10 February 1675/6 compliments Hooke on his observations on thin plates and says 'which I have not scrupled to make use of so far as they were for my purpose'. And on 5 February 1675/6 he wrote to Hooke: 'You have added much several ways, and especially in taking the colours of thin plates into philosophical consideration. If I have seen further it is by standing on the sholders of Giants. But I make no question but you have divers very considerable experiments besides those you have published, and some it's very probable the same with some of those in my late papers. Two at least there are which I know you have observed, the dilatation of the coloured rings by the obliquation of the eye, and the apparition of a black spot at the contact of two convex glasses and at the top of a water bubble; and it's probable there may be more, besides others which I have not made.' This makes his ungenerous lack of reference to Hooke in the *Opticks* all the more deplorable.

Hooke's theory of light is that it is produced by a very rapid vibratory motion, of very small amplitude, in the luminous body, which need not, however, be hot, since rotten flesh and wood give out light, but are cold. The light is carried by an all-pervading transparent homogeneous medium—the ether, as all the classical writers on light called it. In the free ether the 'orbicular pulse' is at right angles to the ray—whether this is to be called a transverse wave is a matter which can be lengthily discussed. Like others, Hooke supposed the disturbance to travel faster in the refracting medium, which, by a kind of Huygens construction—long before Huygens's exposition—became a beam with the 'pulse' oblique to it in the medium. It was a happy guess of Hooke's that the vibrations were transverse, but this obliquity in the medium spoilt it. His theory of colour, to which he attached considerable importance, is obscure; he opined that it was caused by refraction, but says that red is the impression produced by an oblique and confused pulse whose strongest part precedes and whose weakest follows, while blue is the impression produced by a like oblique and confused pulse whose weakest part precedes and whose strongest follows. What he meant by 'confused' and by the weakest and strongest parts I do not know. For him red and blue were the primaries and the other colours were produced from them. This theory of colour is not the happiest part of Hooke's theories, and his ill-tempered attacks on Newton's masterly work with the prism is one of the very few cases where he did not appreciate a great discovery.

A passing reference must be made to Hooke's work on diffraction, nine years subsequent indeed to that of Grimaldi, but almost certainly made in ignorance of Grimaldi's work. Hooke speaks of a new property of light not observed, that he knew of, by any optical writer and Priestley* agrees that Hooke's experiments are so different that it is extremely probable that they were completely original. Hooke describes both the deflexion of the light and the colours produced by it. He compared it to the straying of sound into the quiescent medium; Newton, however, said that it was only a new kind of refraction, caused by a variation in the density of the ether near the edge of the solid body. Newton likewise pointed out, no doubt with some satisfaction, that he had found the phenomenon in Faber's *De Lumine* and that Faber had taken it from Grimaldi. Hooke, however, with his usual acuteness, hit upon the allied phenomenon of diffraction in sound for comparison.

As regards sound, Hooke did not do very much, but what he did was, as usual, an anticipation of much that was done again later. 'He showed an experiment for making musical and other sounds by the help of teeth of brass wheels, which teeth were made of equal bigness for musical sounds, but unequal for vocal sounds',† which is Savart's wheel (date 1820) with an added refinement. He carried out experiments on the vibrations of an inverted glass bell full of water; he stated that the motion of the glass was from oval to circular 'and presently changed into an oval the other way', which he discovered by the rising of water in the glass. 'Sir Christopher Wren coming in said, that the glass would vibrate much stronger, being struck on the edge with a viol-bow.'‡ They then obtained the vibration in six and in eight segments; the four and the eight were octaves and the six and the four were fifths. These observations on bells are usually attributed to Chladni, who carried out his experiments in 1787, using water, just as did Hooke. Hooke also did something in the way of obtaining Chladni figures with flour as the powder, 'it being', says the account, 'manifest by this experiment, that as every different stroke makes a different sound, so the making of a different impression upon the flour gives it as many several motions';§ or, in other words, the flour indicated a different mode of vibration for each note. He also found the frequency of a given musical note with a monochord,‖ a matter which he duly discussed with Samuel Pepys. 'Discoursed with Mr Hooke about the nature of sounds, and he did make me understand the nature of musicall sounds made by strings, mighty prettily; and told me that having come to a certain number of vibrations proper to make any tone, he is able to tell how many strokes a fly makes with her wings, those flies that hum in their flying, by the note that it answers to in musique, during their flying. That, I suppose, is a little too much refined; but his discourse in general of sound was mighty fine.'

* *History of Vision, Light and Colours*, 1772, p. 172. The reference which Priestley there gives to Birch is wrong; it should be Birch, III, 194, which in its turn refers to the *Posthumous Works*, p. 186.
† Birch, IV, 96. See also Waller, xxiii.
‡ Birch, IV, 46.
§ Birch, II, 475.
‖ Birch, I, 446, 451.

Hooke tells us in the *Micrographia* that he had experimented on the propagation of sound through wires over a very considerable distance and adds that it travelled in the wire incomparably swifter than in the air—it actually does travel some fifteen times as fast in an iron wire, say, as in air. He seems to have considered this as a possible method of communication.

Hooke is, of course, one of the great classical microscopists, whose name would live with those of Malpighi; Grew, whom he probably inspired; Swammerdam; and Leeuwenhoek had he left behind only his microscopic observations. His pictures of moulds, of moss, of the nettle, and of insects and their parts, such as the bee's sting and the fly's eye, have been objects of admiration ever since they were executed. The figures of the flea and the louse were long famous. I have a little book called *The Wonders of the Microscope*, published in 1811, which figures a Culpepper microscope of the type made by Harris in about 1780, but is illustrated entirely by Hooke's figures, apparently printed from the original plates. His representation and discussion of the structure of a feather was long the standard one.* His investigation of the microscopic structure of cork led him to discuss the cellular structure of plants, which he was the first to recognize; it is to him that we owe the introduction of the word 'cell' in its biological sense. Perhaps here mention may be made of his work on the life cycle of the gnat, which is a biological investigation of great significance.

As a geologist and early evolutionist Hooke was far in advance of his time. For his contemporaries geology was mainly a subject for fanciful speculations as to the origin and structure of the earth, which made little use of observation and material evidence, but much use of biblical legend and of the teaching of the church in general. For instance, two contemporaries of Hooke were Thomas Burnet and William Whiston: Burnet published in 1681 *Telluris Theoria Sacra*,† which appeared later in English as the *Theory of the Earth*; here he maintained that the earth was like a gigantic egg, the shell of which at the time of the Deluge—the great geological upheaval—was split, with the consequence that internal waters gushed out, and so on: it was the wickedness of mankind that led to the disaster. William Whiston, in his *New Theory of the Earth*, which appeared in 1696, supposed that the rotation of the earth began with the fall of man, and that, as with Burnet, at the time of Noah the central abyss of waters broke through the crust, a comet having something to do with it. Hooke's geology, which is to be found in the *Posthumous Works*, stands out in sharp contrast to these fanciful romances. For him fossils were a definite record of past life, and not mere curious freaks of nature. He made excellent drawings, of which those of the ammonites (figure 15) are typical—W. N. Edwards says that they are not always equalled to-day—and argued forcibly that fossils are either organisms themselves turned to stone or else impressions left by such organisms. The fact that remains of marine creatures were found high above the sea he put down to the upheavals of the earth's surface which have taken place: it should be noted that he used the term 'earthquakes' to denote any

---

* To reproduce here any of the figures mentioned, on the reduced scale that would be necessary, would not do them justice. Reference should be made to the originals.

† There was a second volume of the Latin work in 1689, and of the English in 1690.

serious displacements of the earth's surface. He discounts Noah's flood, so popular with his contemporaries as an explanation of the past life of the globe, pointing out that it cannot have lasted long enough for the production of the shells which are found high on dry land—'besides the quantity and thickness of the Beds of Sand with which they are many times found mixed, do argue that there must needs be a much longer time of the Seas Residence above the same, than so short a space
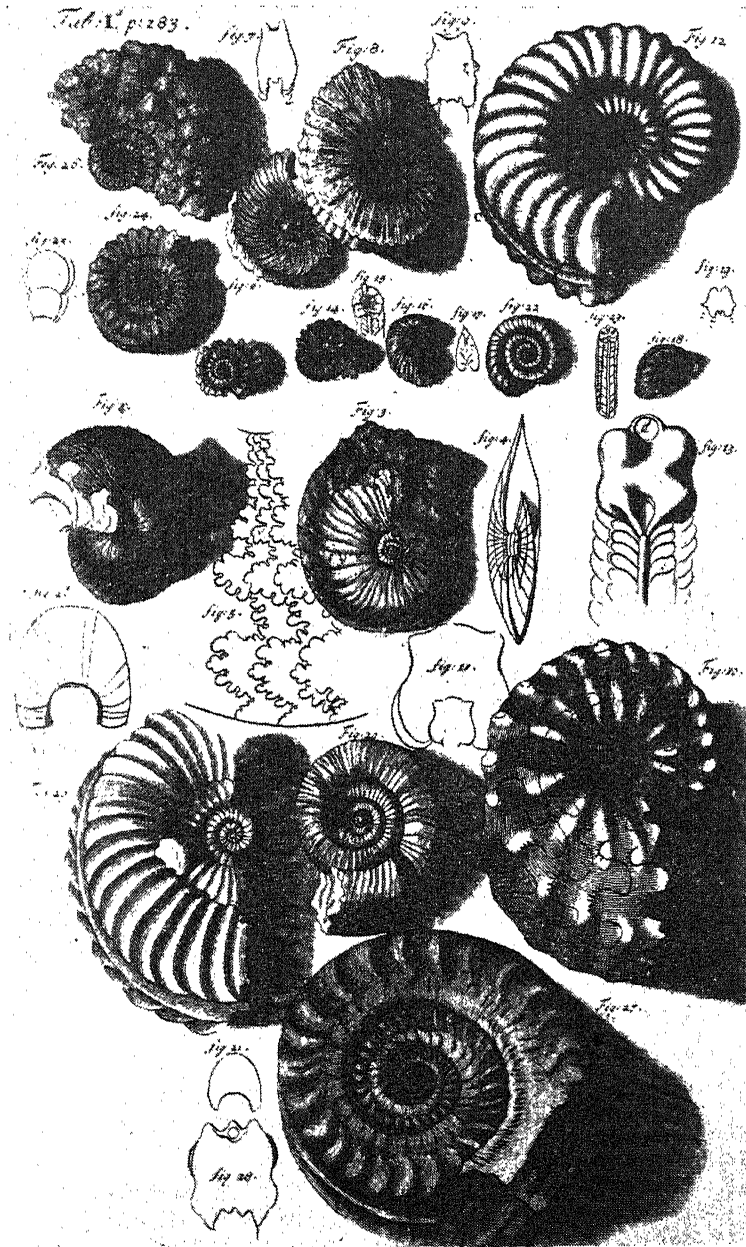

FIGURE 15. Hooke's drawings of ammonites.

can afford'. His general opinion, so much in advance of his time, let him give in his own words, words that are a good example of his beautiful, clear and direct prose: 'I do therefore humbly conceive (tho' some possibly may think there is too much notice taken of such a trivial thing as a rotten Shell, yet) that Men do generally too much slight and pass over without regard these Records of Antiquity which Nature have left as Monuments and Hieroglyphick Characters of preceding Transactions in the like duration or Transactions of the Body of the Earth, which are infinitely more evident and certain tokens than any thing of Antiquity that can be fetched out of Coins or Medals, or any other way yet known, since the best of those ways may be counterfeited or made by Art and Design, as may also Books, Manuscripts and Inscriptions, as all the Learned are now sufficiently satisfied, has often been actually practised; but those Characters [fossil shells] are not to be counterfeited by all the Craft in the World, nor can they be doubted to be, what they appear, by any one that will impartially examine the true appearances of them: And tho' it must be granted, that it is very difficult to read them, and to raise a *Chronology* out of them, and to state the intervalls of the Times wherein such, or such Catastrophies and Mutations have happened; yet 'tis not impossible, but that, by the help of those joined to other means and assistances of Information, much may be done even in that part of Information also.'* I myself well remember as a little boy finding fossil shells on the cliffs at Alum Bay in the Isle of Wight, high above the Needles, a couple of miles from Hooke's birthplace, and I like to think that it was, perhaps, boyhood memories of some such finds that interested Hooke particularly in the matter.

There is no time to refer briefly to more than a small part of his astonishing speculations on cosmological matters. The sun, he says, is solid, encompassed with a bright atmosphere, clouds in which are the sun spots and movements of which give the faculae; all this he argues from observed facts. The light of the sun is due to heat produced by some kind of combustion; sulphur and nitre make a very bright flame, but probably the materials of the sun are much better adapted for producing heat—a conjecture which recent discovery has certainly confirmed! The stars are probably bodies much like the sun. The nearest of them, he deduces from measurement, cannot be nearer than 68,760,000 earth diameters, that is, 35 light days—not far enough, but pretty good for the time. And, he says, other stars are much farther and more powerful telescopes will show still farther and farther stars 'certainly this Material Expansum, part of which we are, must be so great that 'twill infinitely exceed our shallow Conception to imagine'. Again, 'To me indeed the Universe seems to be vastly bigger than 'tis hitherto asserted by any writer.'† Of comets he gives much better pictures than does Hevelius, who spent his whole life observing. He explains the behaviour of the comet's tail by repulsion from the sun, which we know now to exist on account of the pressure of light, while clearly

---

* *Posthumous Works*, p. 411.

† Sir Harold Spencer Jones, the Astronomer Royal, has kindly pointed out to me that already in 1576 Thomas Digges had stated that greater distance was the reason of the relative lack of luminosity of the fainter stars, saying 'the greatest part rest by reason of their wonderfull distance inuisible vnto vs'. See F. J. Johnson, *Astronomical Thought in Renaissance England,* 1937.

recognizing that the body of the comet is attracted by the sun. 'So that though I suppose the attractive power of the Sun...may draw the body towards it, and so bend the motion of the Comet from the straight line, in which it tends, into a kind of curve, whose concave part is towards the sun...yet I conceive that all those parts of the Comet which are thus wrought upon by the other, and changed into another state, and are very much rarified, and produce light, are of a clear contrary nature, and recede from the centre of the Sun.' He was the first to observe and record a spot on Jupiter; he watched it for two hours and observed that the planet rotated on its axis, but neglected to determine the period. He pointed out that the rings of Saturn are sensibly more luminous than the planet. He observed spots on Mars and from his observations concluded that the planet rotated in a period of 12 or 24 hours. The true period is 24 hours. Actually Hooke's drawing of the markings on Mars was used in recent times by Proctor to determine this period with great accuracy, as 24 hr. 27 min. 32·2 sec.! In 1665 he discovered that γ Arietis was actually two stars 8 sec. apart—one of the earliest records of a double star. One of his most careful and systematic observations was the attempt to observe stellar parallax with a fixed vertical telescope, recorded in *An Attempt to Prove the Motion of the Earth*. He selected the star γ Draconis as his object and concluded, from its apparent position at different times of year, that there was a detectable parallax.* It was found, however, that the displacement of the star presented strange features that could not be explained by parallax, and it was as a consequence of very careful observations, which confirmed the accuracy of Hooke's, that Bradley in 1727, nearly sixty years later, discovered aberration, namely, that the apparent displacement was a consequence of the finite velocity of light.
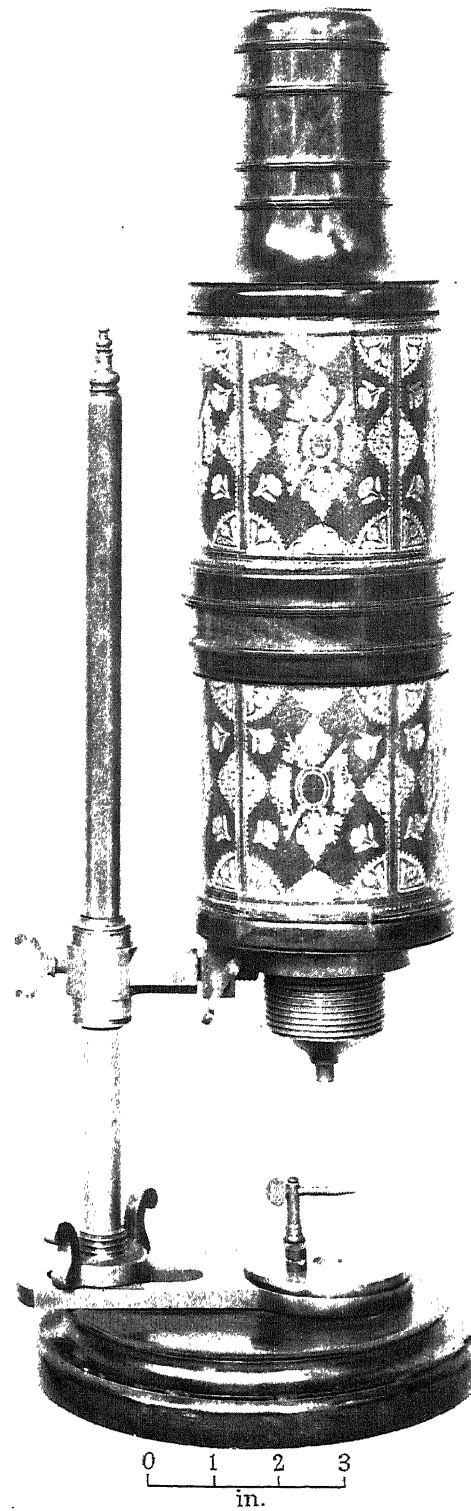
As regards theoretical astronomy, Hooke most clearly knew the principles on which the planetary system was to be explained; he was, I think, the first to state plainly that the problem was one of ordinary dynamics, with no mystical element. In 1674 he said that his system of the world depended upon three suppositions: first, that all celestial bodies whatever had a gravitating power towards their centres, the one for the other; secondly, that all bodies continue in a straight line motion except they are bent aside by some force—'describing a Circle, Ellipsis, or some other more compounded Curve Line'. The third supposition was that the attractive power fell off with the distance, but he did not as yet know the law. Of this law he says: 'it is a notion which if fully prosecuted as it ought to be, will mightily assist the Astronomer to reduce all the Celestial Motions to a certain rule, which I doubt will never be done true without it'. On 6 January 1680, however, Hooke wrote to Newton 'my supposition is that the attraction always is in duplicate portion to the distance from the centre reciprocall', which is the inverse square law in the language of the day. On 17 January he wrote: 'It now remains to know the proprietys of a curve line (not circular nor concentrical) made by a central attractive power which makes the velocities of descent from the tangent line or equal straight motion at all distances in a duplicate proportion to the distances reciprocally taken.' If Hooke could have solved this he would have had the glory of

---

* '...et c'est bien injustment qu'on l'a soupçonné d'avoir ajusté ses observations sur un système adopté d'avance'. Lalande, *Astronomie* (Troisième Édition), Vol. III, p. 81.

a full explanation of the solar system in terms of universal gravitation. He could not solve it, Newton did. But to have enunciated the problem so clearly was a great advance. Surely he did deserve a mention! Newton had it all already, but how was Hooke to realize that?

We now come to consider Hooke's character. The hardest things have been said of him. He has been depicted as melancholy, mistrustful, miserly, penurious, solitary, jealous of all his contemporaries, secretive, claiming all discoveries for himself. Let me admit at once that he was irritable, that he was a man of strong dislikes, and that he particularly disliked being robbed of credit or cheated of money. He was intolerant of fools. He often complained; he often had something to complain of. His mind was so fertile that he was probably often justified in saying that he had made inventions before others. If he sometimes claimed too much credit, he was often denied any credit, as in the passages with Newton. Here he had the, for him, unspeakable chagrin of having before all his fellows attained clear notions on fundamental problems, only to see the greatest of all men of science carry the matter to the final victorious conclusion which was beyond his power. Especially was this the case in the matter of the planetary motions, but his experiments on light also did much to show the way. As I have pointed out, Newton at one time acknowledged that he had learnt much from Hooke's work, but in the end he gave no thanks. The greater man could, as appears from Hooke's letters, so easily have satisfied him.

Yes, he was bitter and bad tempered on many an occasion. He complains that Oldenburg did not enter up in the Minute Book his discourses before the Royal Society: in his Diary he calls him 'a lying dog', 'treacherous and a villain'. Well, there is clear proof that Oldenburg did neglect his work. Pepys enters in his Diary that on 21 February 1665/6 he went with Lord Brouncker to Gresham College and heard a good lecture of Mr Hooke's about the trade of felt-making, 'very pretty'. We have Hooke's notes of the lecture, but there is no record of it in the minutes. The drawing which Hooke did to illustrate these notes, reproduced in plate 12, is one example of his powers as a draughtsman. On 1 March 1664/5 Pepys went again to Gresham College, where Hooke, he says, lectured about the late comet 'among other things, proving very probably that this is the very same Comet that appeared before in the year 1618 and that in such a time probably it will appear again', which, as far as I know, was the first suggestion of a periodic orbit for a comet. Again, no account. Oldenburg does, in general, seem to have done his best to discredit Hooke's work. There is no account of Hooke's book *Lampas* in the *Transactions* edited by Oldenburg. Many instances can be given of the way in which Oldenburg intrigued against Hooke. That the Society took Oldenburg's part against him on the occasion of their open dispute was a great blow to Hooke. To Hooke's precise accusations Oldenburg replied: 'The publisher of this tract intends to take another opportunity of justifying himself against the aspersions and calumnies of an immoral postcript put to a book called *Lampas*, published by R. Hooke, till which time, it is hoped, the candid reader will suspend his judgment', but never said any more. It may be noted that Newton liked Oldenburg not much better than Hooke did, in spite of the obsequiousness shown by Oldenburg to him,

Hooke's drawing of the process of felt-making. The lecture which Hooke gave on felt-making is referred to in Pepys' *Diary*.

An original Hooke microscope, now in the Science Museum. (Crown copyright. From the Court Collection in the Science Museum. Reproduced by kind permission of the Science Museum.)

Church at Willen in Buckinghamshire designed by Robert Hooke for Dr Busby.
(From a water-colour by A. E. Richardson, R.A.)

writing, for instance, 'moved by other reasons to decline, as much as Mr Oldenburg's importunity and ways to engage me in dispute will permit'.

In general, it can be shown that Hooke had much to bear from Oldenburg. The Society, too, did not treat him very well. As Curator he was always 'ordered' to 'try the experiment', 'try this by himself at home' or to 'bring in an account in writing'. Although—or because—for long periods he brought in a new experiment at practically every meeting, for little and very irregular pay, on 14 November 1670 'it was resolved that Mr Hooke be summoned to attend the next meeting of Council to receive their rebuke for the neglect of his office'.

Yes, Hooke was irritable and vain—he thought that he had performed great things in science, which is no sin, but he sometimes said so, which is. The fortunate man gets someone else to say it for him, as Francis Bacon points out in his Essay *Of Friendship*. But he was sick, very sick; an ill man day in and day out. Only occasionally, however, was he sorry for himself or too ill to work; only occasionally do we have such an entry as 'Ill all day, miserere mei deus'—God have pity on me! If he ate anything that 'agreed' he noted it in his diary. Terrible headaches, vomiting, giddiness, sleeplessness, fearful dreams all vexed him. He was perpetually administering medicaments of various kinds to himself, in the hope of relief; whether on the whole this did more harm than good, who shall say?

Yes, Hooke was irritable, especially when he thought himself slighted. He was impatient with fools. He was vexed with Sir John Cutler for promising him a stipend and then, after being much praised for his generosity, not paying it. But mark one or two strange things. Although Thomas Molyneux said that he was 'the most ill-natured, conceited man in the world, hated and dispised by most of the Royal Society'—why not by all?—he was unanimously asked to be Secretary when Oldenburg died. He was on good terms with most of the first-class intellects of the time—he dined with Evelyn and stayed for weeks at his house, was invited by Sydenham (who liked Oldenburg as little as Hooke did) to stay with him for six weeks, was on intimate terms with Christopher Wren (there is an entry of his buying a hobby horse for Wren's son), was very friendly with John Aubrey, who wrote of him in the highest terms, and was constantly in the company of Haak, Abraham Hill, Daniel Whistler and others. Busby, his schoolmaster kept up friendship with him all his life; in 1679 Hooke designed a church for him (plate 14). As for Tompion, the famous clockmaker, he occasionally enters him as 'a slug', 'a rascall', but they worked together for years, and, as supreme craftsman and supreme inventor, no doubt knew one another's worth—lost their tempers with one another and met next time as if nothing had happened.

We are told that Hooke was mean and miserly. In particular Ned Ward—'the filthy and facetious Ned Ward'—describes Hooke as a member of the Split-Farthing Club. My belief (for which I have no evidence except the characters of the two men) is that Ward, who was an innkeeper, had attempted to swindle Hooke and come off worse, for from the diary edited by Robinson and Adams it is clear that Hooke was, anyhow for the years that it covers, a sociable and free-spending man, a frequenter of coffee houses who smoked and drank with his friends. He also entertained them with wine at his lodgings and bought brandy by the gallon and wine by the dozen.

An entry like 'Drunk. Promised to pay whatever I signed to be paid' shows his usual directness, but is hard to reconcile with the miserable recluse that he is painted as by ill-wishers. He also drank tea, which cost 25 shillings a pound, at a time when a shilling was a purchasing sum. It is quite possible that Hooke was a man who did not mind spending freely on himself and his friends but disliked wasting money or being cheated—a character that always puzzles the English but not the Scots. He spent large sums on books. Yes, he was an irritable sick man, but he had his friends—and some of them were the most outstanding men of the time— and they seem to have stuck to him. And, at his funeral, anyway, the Royal Society behaved handsomely. John Ward, author of the *Lives of the Professors of Gresham College*, says well: 'Had he been more steady in his pursuits, and perfected one discovery before he entered upon another, he might perhaps in some cases have done greater service to the public, and prevented what often gave him uneasiness, the fear of losing credit of them by others, who built upon his foundation.' Always it must be remembered that he was without influence or position, other than that which he won for himself the hard way. He had no family, no profession of power, no institution behind him. It was not till 1691 that he was awarded his doctorate. In general, it seems to be just to say that while irritable, suspicious and prone to sudden anger under provocation, often severe, he bore no grudge to great men, however they treated him, and was, in large matters and in his devotion to science, a brave, and in some respects a noble, character.

You may think that, carried away by youthful enthusiasm, I have exaggerated the genius of Hooke, that, when I say that I have not even referred to half his achievements, I am using the language of political protestation. Let me, then, cite a few of the judgements of men of enduring reputation in science who have both studied Hooke's works and pronounced upon them. Thomas Young, writing of the interference of light says: 'This, I assert, is a most powerful argument in favour of the theory which I had before revived: there was nothing that could have led to it in any author with whom I am acquainted, except some imperfect hints in those inexhaustible but neglected mines of nascent inventions, the works of the great Dr Robert Hooke....' Lalande, the great astronomer, referring to the *Micrographia*, *Cutlerian Lectures* and *Posthumous Works* says: 'Dans ces ouvrages on trouve les idées de la plupart de nos instruments modernes: c'était le Newton de la mécanique.' Another famous astronomer Mädler, says: 'Denn dass Hooke zu den scharf-sinnigsten Denkern gehört, die jemals gelebt, kann nicht verkannt werden.' Lasswitz, who wrote the standard history of atomism, says of Hooke that he is 'eine der eigentümlichsten Erscheinungen unter den hervorragenden Forschern seiner Zeit. An allen wichtigeren Fragen der Astronomie, Mechanik und Physik beteiligt...macht er mit rastloser Arbeitskraft eine grosse Anzahl wertvoller Erfindungen und Entdeckungen', and devotes a chapter to his contributions to the theory of atomic vibrations. Turning to geology our own Sir Archibald Geikie, a cautious Scot, refers to him as 'one of the most brilliant, ingenious and versatile intellects of the seventeenth century', and, considering his geological work, speaks of 'his remarkable powers of acute observation and sagacious reflection'. Finally, J. F. W. Herschel calls him 'The great contemporary, and almost the worthy rival,

of Newton'. I have chosen an English physicist; a French, a German and a British astronomer; a German philosopher and a British geologist, all of first rank—I could, I assure you, without difficulty have added lesser men to the number. I think that I can claim that all those who have gone direct to Hooke have conceived the highest admiration for his astonishing industry, his whole-hearted devotion to science, his inventiveness, his ingenuity, his fertility and his brilliant theoretical insight. He rendered, with unflagging zeal, services to our Society which have not received the deserved acknowledgement. In his time the Society yielded him scant recompense, and, at any rate in his youthful prime, a grudging recognition. I have tried, while not concealing his faults, to win your sympathy for a much tried man and your approbation, your respect for his astonishing accomplishments. Admire Robert Hooke; he is worthy of your admiration.

A study of the component glycerides of cow and buffalo milk fats with reference to the possible mechanism of their production during lactation

By K. T. ACHAYA AND T. P. HILDITCH, F.R.S.

Department of Industrial Chemistry, University of Liverpool

The component glycerides of a cow and a buffalo milk fat have been studied quantitatively after resolving them by intensive crystallization from acetone or ether at temperatures between -50 and $0°$ C into groups of glycerides which, although still complicated in fatty acid composition, contain much simplified mixtures of mixed glycerides as compared with the original fats. From the component acids of each separated glyceride group and of the trisaturated glycerides present therein, it has been possible to produce an approximately quantitative statement of the many individual mixed glycerides present in the original fats, and a more closely quantitative statement of the various glyceride categories (trisaturated, mono- or di-unsaturated, mono- or di-palmito, mono- or di-oleo, etc.) present.

The bearing of the data so obtained on the question of the mechanism of the production of the characteristic short-chain acyl glycerides of ruminant milk fats is discussed. Evidence is collected which favours the view that the main source of milk fats is glycerides in the blood entering the lactating gland, the oleo- (or other unsaturated) groups in which are transformed in greater or less degree into acyl groups with shorter carbon chains (down to butyric, C_4). It is shown that the alternative hypothesis of direct synthesis in the mammary gland of short-chain glycerides from carbohydrate or acetate is inconsistent with the quantitative composition of the component acids of milk fats of a range of non-ruminant as well as ruminant mammals. On the other hand, it is considered that the maximal production of short-chain acyl glycerides observed in milk fats of the ruminant animals is clearly connected with the presence of unusual amounts of acetic and associated lower fatty acids in their blood, and that this causes conditions (pH range or other factor) in the lactating gland which favour the action of the enzyme system responsible for the conversion of oleo- into short-chain glycerides.

INTRODUCTION

The large number of fatty acids combined as glycerides in ruminant milk fats has rendered the determination of the many mixed triglycerides present a matter of

great difficulty. Systematic crystallization of solid fats which contain a simple mixture of combined fatty acids leads to their resolution into groups of glycerides which, although still usually mixtures of mixed triglycerides, are sufficiently simplified to permit the proportions of their individual constituents to be derived from the total proportions of the fatty acids present in each group with, in appropriate cases, additional data as to the amount and nature of any trisaturated glycerides present (Hilditch 1947, pp. 520–526). A study of an English cow milk fat (Hilditch & Paul 1940) by this procedure (systematic crystallization from acetone at 0 and at 20° C) resulted in its separation into three groups, each of which was still too complex to permit a precise statement of the constituent glycerides of the whole fat to be reached. Nevertheless, the results showed that palmitic acid was present in about 70 % of the triglyceride molecules (i.e. in nearly as high a proportion of the glycerides as in the corresponding body fat) and that about 22 to 30 % of the milk glycerides contained one palmitic and one oleic radical with one radical of a short-chain (C_4, C_6, C_8 or C_{10}) fatty acid; whilst nearly 40 % was made up of oleopalmitostearins, palmitodiolein, triglycerides with one oleic, one stearic, and one short-chain acyl group, and of triglycerides with two oleic groups and one short-chain acyl group. Trisaturated glycerides formed 19 % of the fat, chiefly glycerides with one palmitic, one stearic and one short-chain acyl group, and glycerides with one palmitic and two short-chain acyl groups. No sign of the presence of tributyrin or other simple triglyceride was observed; tri-unsaturated glycerides did not amount to more than 7 % of the milk fat and consisted almost certainly of glycerides with two oleic and one other unsaturated acyl group.

Later it has been found that by employing a range of temperatures down to − 60° C the glycerides of most natural fats (including those liquid at the ordinary temperature) can be resolved by systematic crystallization from appropriate solvents into sufficiently simple mixtures to enable the glyceride constituents of the latter to be deduced from their component acids. In this way the component glycerides of a number of liquid vegetable seed fats have been approximately determined (e.g. Hilditch & Maddison 1940, 1941; Hilditch, Laurent & Meara 1947; Hilditch, Holmberg & Meara 1947; Dunn, Hilditch & Riley 1948), whilst the method has also been successfully applied to marine animal oils, which contain a much more complex mixture of component acids (Hilditch & Maddison 1942, 1948; Bjarnason & Meara 1944; Hilditch & Pathak 1947). A partial application of this low-temperature crystallization to a milk fat was made by Jack & Henderson (1945) and Jack, Henderson & Hinshaw (1946) who cooled a solution of cow milk fat in light petroleum to progressively lowered temperatures and removed the glycerides which separated successively at −7, −13, −23 and −53° C. These differed in character as shown by their respective iodine values (8·3, 22·1, 23·5, 31·2), whilst the material left in solution at −53° C had an iodine value of 58·4. Jack et al. determined the component acids and the proportion of trisaturated glycerides in each of the five groups, and concluded from their data that the fatty acids in milk fat tend to be distributed among the glycerides as widely as possible.

The results of the studies of Hilditch & Paul (1940) and of Jack et al. (1945, 1946), in which a certain degree of separation of milk fat mixed glycerides was thus

effected by the crystallization procedures employed, on the one hand confirmed earlier conclusions on the specific glyceride structure of milk fats which had been reached by previous workers (e.g. Hilditch & Sleightholme 1930, 1931; Dhingra 1933, 1934; Achaya & Banerjee, 1946) from consideration of the proportions of trisaturated glycerides and of their component acids, and on the other hand indicated that fuller application of the later, more thorough low-temperature crystallization techniques was desirable. We have therefore submitted specimens of Indian cow and buffalo milk fats to an elaborate scheme of crystallization from acetone or from ether at temperatures down to $-50°$ C, and have obtained data which, although still not precise, give the most complete picture so far obtainable of the many mixed triglycerides which go to constitute a typical bovine milk fat. These results have enabled us to make some further observations as to the possible origin of the short-chain acyl groups (from butyric to dodecanoic) which are specifically characteristic of this group of fats.

The specimens of cow and buffalo milk fats (ghees) studied were conveyed from the Imperial Dairy Institute, Bangalore, in September 1946 in carefully sealed, full air-tight tins which were kept unopened in cold storage until the contents were required for use. Both fats were found to be in excellent condition with negligible free fatty acid contents and good fresh aroma. All the animals had been fed during the milking period on daily rations consisting approximately of Ragi straw 3 lb., Guinea grass 30 lb., Rhodes grass 25 lb., groundnut cake 5½ lb., gram husk 4½ lb., crushed gram 3 lb., wheat bran 8 lb., and salt 1 oz. The cows were of the Gir, Sindhi, Tharparkar and Hariana breeds, and the buffaloes of the Murrah and Surti breeds.

The general analytical characteristics of the two milk fats are given in table 1.

TABLE 1. GENERAL ANALYTICAL CHARACTERISTICS OF THE INDIAN
COW AND BUFFALO MILK FATS

	cow milk fat	buffalo milk fat
iodine value	34·3	30·2
Reichert value	25·4	28·8
Polenske value	2·5	1·9
saponification value	226·0	226·7
saponification equivalent	247·1	247·5
free fatty acid (as oleic) (%)	0·3	0·1
unsaponifiable matter (%)	0·2	0·3
refractive index (40° C)	1·4530	1·4539
carotene (mg./g.)	18·1	nil
vitamin A (i.u./g.)	8·3	16·5

METHODS

Low-temperature crystallization of the milk fats

Several preliminary trials on the Indian cow milk fat were undertaken in order to define the general range of solvents, concentration of solutions and temperatures of crystallization likely to be most suitable. The separations illustrated in figure 1 (buffalo milk fat) and figure 2 (cow milk fat) were subsequently effected. (In the

crystallization diagrams all weights are corrected to the original amount taken; the solvent employed is indicated by the letters E (ether) or A (acetone); and the number before the letters E or A (e.g. 5 E or 10 A) refers to the concentration of the solution in ml. of solvent per g. of fat.) Briefly, it was found best in the preliminary

790·6 g., i.v. 30·2

5 E, −30° C

600·2 g., i v. 25·8   190·4 g., i.v 56·2

7 E, −30° C   10 A, −50° C

526·5 g., i.v. 22·6   73·7 g., i.v. 47·8   164·5 g., i.v. 45·8   25·9 g., i.v. 61·0

10 A, −40° C

194·3 g., i.v. 43·4   43·9 g., i.v. 55·0

10 A, −40° C

174·5 g., i.v. 43·1   19·8 g., i.v. 52·8

8 E, −20° C

364·4 g., i.v. 16·9   162·1 g., i.v. 33·5

10 E, −20° C

298·3 g., i.v. 13·5   66·1 g., i.v. 33·5   10 A, −30° C

184·5 g., i.v. 30·7   43·7 g., i.v. 53·3

10 E, 0° C   10 A, −30° C

73·8 g., i.v. 4·6   160·2 g., i.v. 28·6   24·3 g., i.v. 49·1

A   224·5 g., i.v. 16·2   20 A, −10° C   F

98·5 g., i.v. 24·6   61·7 g., i.v. 34·9

10 A, 0° C

69·5 g., i.v. 23·2   29·0 g., i.v. 27·1

20 E, −20° C   10 A, −10° C

144·2 g., i.v. 13·0   80·3 g., i.v. 21·1   17·6 g., i.v. 38·7   156·9 g., i.v. 44·6
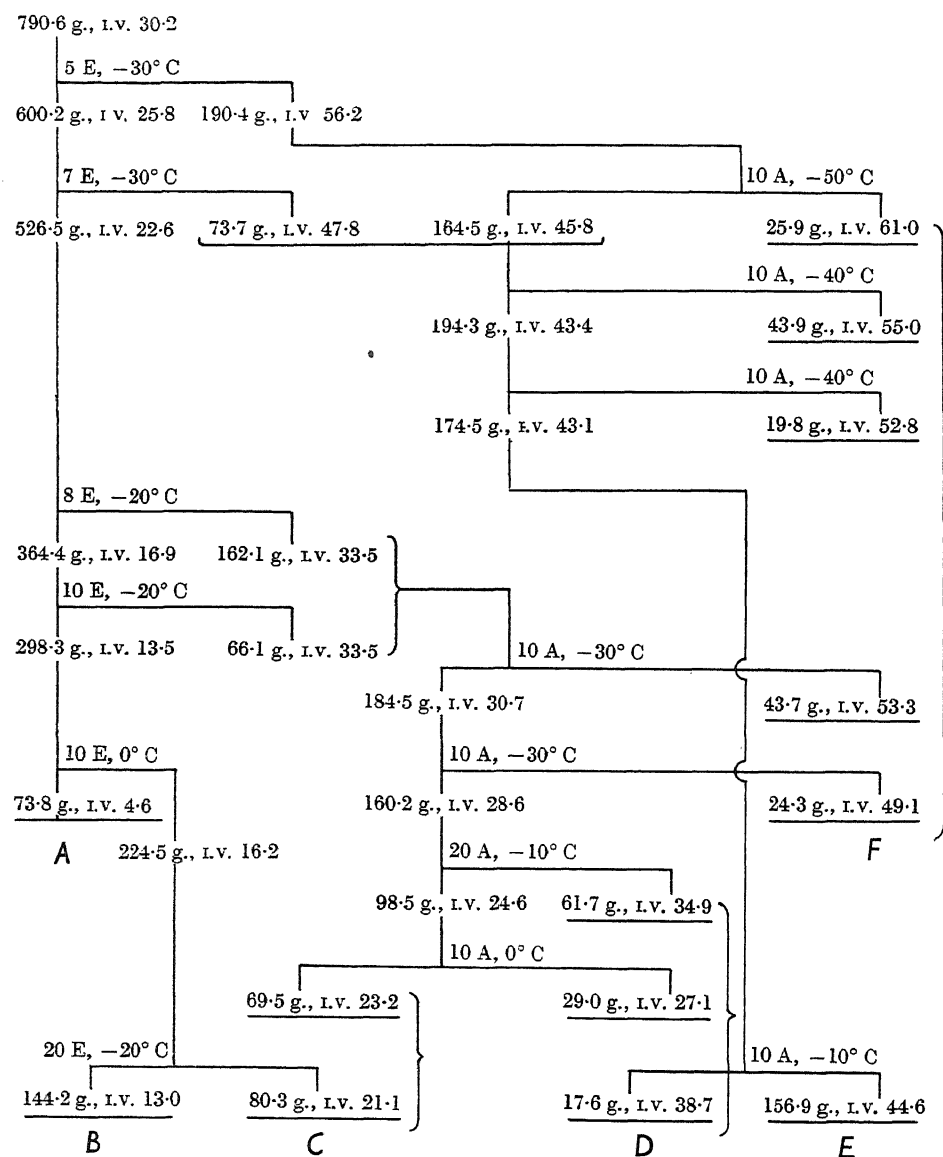
B   C   D   E

Figure 1. Crystallization chart: buffalo milk fat.

stages to crystallize the milk fats (or to recrystallize the fats deposited in a previous operation) from fairly concentrated (5 or 7 ml./g. fat) solutions of the fats in ether at a moderately low temperature (− 30 to − 20° C). Further crystallizations of the deposited, relatively saturated glycerides (e.g. iodine value 15 to 20) from more dilute solutions in ether at about − 20° C and finally at 0° C led to the isolation of

a fraction of low iodine value (*c.* 5) and of others of iodine value *c.* 13 and *c.* 20. On the other side, the fractions left in solution in ether in the earlier crystallizations were more unsaturated, with iodine values of the order of 60 to 40; these were further separated by employing acetone (usually 10 ml./g. fat) as solvent and
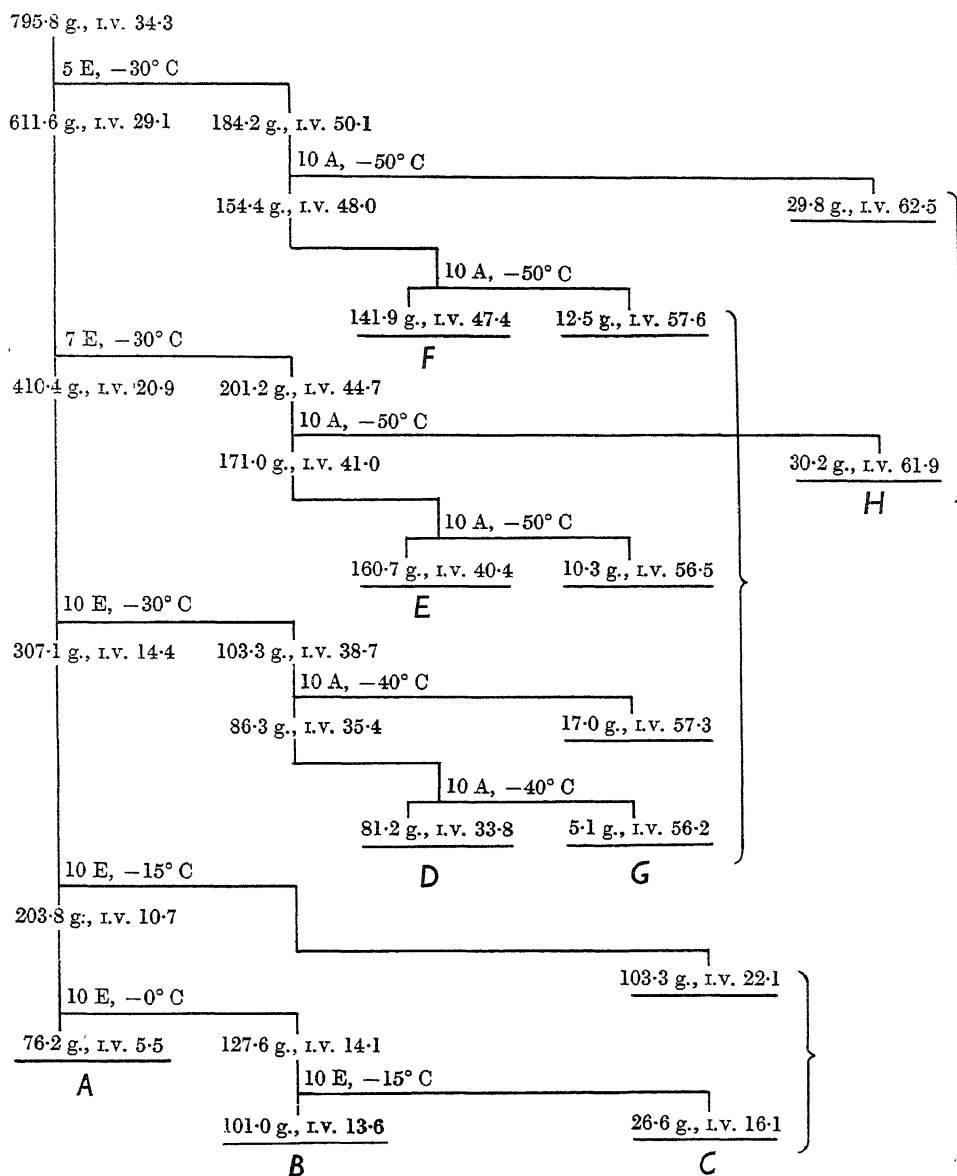

FIGURE 2. Crystallization chart: cow milk fat

operating at temperatures of − 50, − 40 or − 30° C. At various points subfractions of similar iodine value were united for further crystallizations. The influence of more soluble glyceride solutes upon more sparingly soluble glycerides ('mutual solubility effect') renders the separations more difficult and, in practice, never

complete; but it will be seen later that the procedure adopted has considerably simplified the mixture of glycerides present in any of the final fractions.

In all, fourteen operations were required to separate the buffalo milk fat into six final groups; whilst eight final fractions were obtained from the cow milk fat as a result of twelve separate crystallizations. In the final assembling of fractions for further detailed examination, subfractions of closely similar iodine value were in several instances combined together, as indicated in figures 1 and 2. The proportions and general characteristics of the fractions and of the whole fats are shown in table 2.

TABLE 2. GENERAL SUMMARY OF GLYCERIDE FRACTIONS OBTAINED BY CRYSTALLIZATION OF INDIAN BUFFALO AND COW MILK FATS

	buffalo milk fat					cow milk fat				
fraction	wt. (g.)	% (wt.)	% (mol.)	iodine value	sapon. equiv.	wt. (g.)	% (wt.)	% (mol.)	iodine value	sapon. equiv.
A	73·8	9·3	8·5	4·6	272·0	76·2	9·6	8·8	5·5	268·4
B	144·2	18·2	18·4	13·0	245·4	101·0	12·7	12·7	13·6	246·8
C	149·8	19·0	18·6	22·0	253·7	129·9	16·3	16·3	20·8	246·2
D	108·2	13·7	13·7	32·7	248·0	81·2	10·2	10·1	34·2	250·8
E	156·9	19·8	19·5	43·4	252·9	160·7	20·2	20·2	40·4	246·8
F	157·7	20·0	21·3	52·8	232·2	142·0	17·8	17·7	47·3	248·0
G	—	—	—	—	—	44·8	5·6	6·1	56·8	228·5
H	—	—	—	—	—	60·0	7·6	8·1	60·5	228·4
(original fat	790·6	100·0	100·0	30·2	247·5	795·8	100·0	100·0	34·3	247·1)

Detailed investigation of each fraction

This consisted in the determination (a) of the component acids in each of the fourteen fractions, (b) of the amount of trisaturated (fully saturated) glycerides in all but two of the fractions, and (c) of the component acids present in each of the trisaturated glyceride fractions isolated. In addition, as a check on the probable order of accuracy of these analyses, similar determinations of component fatty acids and trisaturated glycerides were carried out on the original milk fats. These determinations were made by the procedures which have been in general use in our laboratory for a considerable time, but certain fresh modifications which were found convenient may be briefly described.

Detection of trisaturated glycerides

This was effected by the method of Hilditch & Lea (1927) as described by Hilditch (1947, pp. 513–516), with the following modifications which were found to be advantageous:

(i) Powdered potassium permanganate was added to the heated solution of the fat in acetone until the solution retained a permanent purple colour after prolonged refluxing. At this stage vigorous refluxing was continued for several hours, the cake of precipitated manganese oxides being frequently broken up and well stirred during this period.

(ii) The removal of the acidic products (nonoic and hexoic acids and azelao-glycerides) of the disruptive oxidation of the unsaturated acyl glycerides constitutes the main difficulty of the process owing to the emulsifying properties of the semi-acid azelao-glycerides. We have found the following procedure an improvement upon those hitherto described. The solution in ether of the reaction products is first cautiously washed (not shaken) several times with a 10 % aqueous solution of potassium carbonate, the washings being later repeated with gentle agitation. A similar series of washings with N/10 aqueous caustic potash, and finally with N/2 aqueous caustic potash, usually suffices to complete the process, apart from final washings with distilled water to remove all traces of alkali and alkali salts. This procedure gave a minimum of trouble with emulsification and in general led to the isolation of fully saturated glycerides with very low acid values and residual iodine values.

The order of accuracy in trisaturated glyceride determinations by this method is probably somewhat less than in the determination of component fatty acids, and in our experience the results tend to be somewhat lower than the truth. When the proportion of trisaturated glycerides is not very large, analytical error due to this cause will not vitiate the final conclusions to any serious extent. When, as in the most saturated glyceride groups *A* and *B*, there are very large proportions of trisaturated glycerides present, the divergence may be more serious; but in such cases the components may be calculated from the determined proportions of unsaturated acids as a mixture of mono-unsaturated disaturated and trisaturated glycerides, since it can be taken for granted that here di-unsaturated-monosaturated glycerides will not be present. This is well illustrated by groups *A* and *B* from the buffalo milk fat:

A trisaturated glycerides found 82·7 %, minimum (from component acid data) 84·5 %.

B trisaturated glycerides found 58·2 %, minimum (from component acid data) 58·8 %.

Determination of component fatty acids

The methods described by Hilditch (1947, pp. 467–489) for the determination of the component acids of milk fats were used throughout, with the exception that the fatty acids non-volatile in steam were not resolved into groups of mainly saturated and mainly unsaturated acids by the lead salt separation formerly employed; instead, it was found more convenient to crystallize them from 10 % solution in ether at − 40° C. This usually gave an excellent separation of saturated fatty acids (C_{14} to C_{20}) accompanied by very little oleic acid, and left in solution oleic and the minor unsaturated acids with small proportions of saturated (C_{10} to C_{16}) acids. Each group of higher fatty acids was converted into methyl esters which were submitted to fractionation in the usual manner at 0·2 mm. pressure through an electrically heated and packed column. In fractions of comparatively small weight the quantity of steam-volatile acids was too small for convenient fractionation in the usual apparatus, and in such instances a small semi-micro-scale fractionation apparatus was successfully employed. Moreover, the available amounts of some of the trisaturated glycerides were insufficiently large to permit

TABLE 3. DETERMINATION OF THE COMPONENT ACIDS OF THE
INDIAN BUFFALO MILK FAT

I. Preliminary separation of fatty acids

fraction	description	wt. (g.)	%	iodine value	sapon. equiv.
V	acids volatile in steam	9·51	7·9	2·6	102·7
A	acids non-volatile in steam insoluble in ether at −40° C	57·34	48·0	2·1	261·3
B	acids non-volatile in steam, soluble in ether at −40° C	52·71	44·1	72·1	261·7

II. Fractionation data for groups V, A, B.

(i) Volatile acids V (semi-microfractional distillation)

fraction	distillation pressure	column head temp. (° C)	wt. (g.)	sapon. equiv.	iodine value
V1	atmospheric	*	2·695	*	—
V2	atmospheric	168	0·871	90·9	—
V3	atmospheric	170	0·775	91·0	—
V4	atmospheric	174	0·665	91·3	—
V5	atmospheric	182	0·936	94·5	—
V6	atmospheric	186	0·682	100·0	—
V7	atmospheric	196	0·672	113·5	—
V8	atmospheric	206	0·689	125·2	1·2
V9	15 mm.	174	0·417	150·3	4·8
V10	15 mm.	residue	1·104	196·7	19·7

* Acid in ether-extracted aqueous liquor, recovered ether and first runnings (b.p. 110 to 164° C) taken as butyric acid.

(ii) Methyl esters of acids A and B fractionated at 0·2 mm. pressure

	methyl esters of acids A				methyl esters of acids B			
fraction	column head temp. (° C)	wt. (g.)	sapon. equiv.	iodine value	fraction	column head temp.(° C)	wt. (g.)'	sapon. equiv.	iodine value
A1	96	0·75	244·7	0·9	B1	82	0·99	192·5	6·6
A2	101	2·78	249·5	0·3	B2	88	1·09	205·2	4·5
A3	107	3·44	253·7	0·3	B3	101	2·21	221·0	7·5
A4	113	3·65	256·4	0·1	B4	107	2·46	237·3	9·4
A5	118	4·42	264·6	0·2	B5	110	2·39	238·5	7·9
A6	120	1·92	265·6	0·3	B6	120	2·50	248·1	16·8
A7	127	5·05	268·9	0·2	B7	132	3·56	263·6	40·7
A8	128	5·38	269·9	0·1	B8	138	3·82	278·7	68·4
A9	130	6·91	269·7	0·6	B9	140	4·55	286·1	79·9
A10	148	8·77	282·8	2·9	B10	140	3·93	287·3	85·4
A11	152	5·66	296·1	4·5	B11	142	5·08	294·6	89·2
A12	152	9·17	298·7	5·0	B12	143	5·95	293·8	90·2
A13	residue	1·62	356·2	7·9	B13	145	5·99	294·4	90·4
					B14	147	4·08	292·9	91·3
					B15	148	2·64	300·2	100·4
					B16	residue	1·70	399·4*	103·1

* B16 esters freed from unsaponifiable matter, sapon. equiv. 320·4.

III. Component acids of acid groups V, A, B, and of whole milk fat (% wt. and % mol.).

	volatile acids 'V' 7·9% % (wt.)	non-volatile acids		acids of whole milk fat (excluding unsaponifiable matter)	
		'A' 48·0% % (wt.)	'B' 44·1% % (wt.)	% (wt.)	% (mol.)
Proportion of total acids ...					
saturated acids					
butyric	60·4	—	—	4·8	12·6
caproic	21·0	—	—	1·7	3·3
caprylic	6·2	—	—	0·5	0·8
capric	6·1	—	1·8	1·3	1·7
lauric	3·4	—	5·7	2·8	3·3
myristic	—	12·8	12·1	11·5	11·7
palmitic	—	52·8	8·1	29·0	26·3
stearic	—	29·1	—	14·0	11·5
arachidic(?)	—	2·9	—	1·4	1·1
unsaturated acids:					
decenoic	—	—	0·1	trace	trace
dodecenoic	—	—	0·3	0·1	0·2
tetradecenoic	—	—	1·7	0·8	0·8
hexadecenoic	—	—	8·6	3·8	3·5
oleic	2·9	2·4	54·6	25·5	21·0
octadecadienoic	—	—	3·7	1·6	1·3
unsaturated C_{20} to C_{22}	—	—	2·6	1·2	0·9
unsaponifiable	—	—	0·7	—	—

separated removal of the steam-volatile (lower saturated) acids. In these cases the trisaturated glycerides were directly converted into methyl esters by 'methanolysis' (Haller 1906), which were then fractionated by the method described by Smith & Dastur (1938).

The data recorded in the following pages of course show only the final results of the thirty component acid analyses and fourteen trisaturated glyceride determinations which have been necessary in the course of the present study. In order to illustrate the procedure which has now been followed, however, the detailed data relevant to the determination of the component acids of the original buffalo milk fat are reproduced in table 3. Similar data and calculations have naturally been requisite in regard to the analyses of each of the separate milk-fat fractions and of their trisaturated components.

RESULTS

The final analytical data on which interpretation of the glyceride structure of the milk fats depends are collected in tables 4a, 4b, 5a and 5b.

TABLE 4a. INDIAN BUFFALO MILK FAT: COMPONENT ACIDS (% MOL.) IN GLYCERIDE GROUPS A TO F AND THE CORRESPONDING TRISATURATED GLYCERIDES

	A	B	C	D	E	F
%(mol.) of whole fat ...	8·5	18·4	18·6	13·7	19·5	21·3
component acids:						
butyric	0·3	11·8	11·1	10·5	11·1	20·5
caproic	—	3·4	3·6	3·9	3·9	6·1
caprylic	0·2	0·8	0·6	—	1·2	1·5
capric	0·3	1·2	1·8	2·6	1·3	2·6
lauric	—	2·5	1·7	3·3	2·9	3·4
myristic	14·3	10·7	11·6	13·1	10·6	8·8
palmitic	46·8	34·4	34·8	24·3	20·9	15·3
stearic	29·4	19·3	11·6	8·9	5·9	2·5
arachidic	3·5	2·1	1·6	1·8	1·1	—
decenoic	—	—	trace	—	trace	0·1
dodecenoic	—	0·1	0·1	0·1	0·3	0·2
tetradecenoic	0·1	0·4	0·6	0·7	1·4	1·0
hexadecenoic	1·4	0·5	0·8	1·9	2·8	2·7
oleic	3·7	12·7	18·2	24·4	33·6	32·7
octadecadienoic	—	—	0·2	0·8	—	—
unsaturated C_{20} to $_{22}$	—	0·1	1·7	3·7	3·0	2·6
trisaturated glycerides:	84·5	58·8	43·2	44·3	23·6	13·9
% (mol.) of group:						
component acids:						
butyric	0·3	16·4	17·5	16·8	19·3	24·7
caproic	—	5·1	5·7	6·3	5·4	9·3
caprylic	—	1·1	0·9	0·8	2·9	4·3
capric	—	4·5	4·3	2·3	4·2	8·9
lauric	0·3	1·5	5·5	7·5	13·4	14·3
myristic	17·6	10·9	12·9	20·8	15·6	10·8
palmitic	48·7	38·7	38·5	31·5	27·2	18·5
stearic	29·3	19·9	12·6	10·6	9·7	6·2
arachidic	3·8	1·9	2·1	3·4	2·3	3·0

TABLE 4b. INDIAN COW MILK FAT: COMPONENT ACIDS (% MOL.) IN GLYCERIDE GROUPS A TO H AND THE CORRESPONDING TRISATURATED GLYCERIDES

	A	B	C	D	E	F	G	H
% (mol.) of whole fat ...	8·8	12·7	16·3	10·1	20·2	17·7	6·1	8·1
component acids:								
butyric	1·0	9·3	10·0	8·3	8·4	9·2	17·6	20·6
caproic	0·2	4·5	5·1	3·5	3·5	4·8	8·4	7·5
caprylic	—	1·2	1·7	1·9	2·2	2·9	1·8	3·4
capric	0·2	2·4	1·3	2·3	2·7	2·9	4·5	3·5
lauric	1·0	2·3	2·9	3·6	3·3	2·7	2·2	3·1
myristic	15·5	8·4	9·3	12·9	10·9	10·4	7·7	6·9
palmitic	45·3	41·9	34·5	24·3	23·5	17·4	11·2	6·9
stearic	25·4	15·6	13·2	9·3	6·3	4·8	1·3	0·4
arachidic	4·7	1·1	1·2	1·1	0·4	0·8	—	0·3
decenoic	—	0·1	0·1	0·1	0·2	0·2	0·2	0·2
dodecenoic	—	0·2	0·3	0·2	0·4	0·4	0·3	0·6
tetradecenoic	0·1	0·8	1·4	0·9	1·4	1·5	1·7	2·2
hexadecenoic	1·5	1·7	3·2	1·2	5·9	3·1	4·5	4·6
oleic	5·1	10·1	14·6	29·0	29·2	35·8	32·9	33·5
octadecadienoic	—	—	—	0·8	1·2	1·7	3·4	3·9
unsaturated C_{20} to C_{22}	—	0·4	1·2	0·6	0·5	1·4	2·3	2·4
trisaturated glycerides: % (mol.) of group:	81·2	60·3	47·7	24·4	22·5	13·6	22·5*	22·5*
butyric	1·2	8·5	16·4	12·3	20·6	18·9 (not determined)		
caproic	—	3·0	6·9	6·3	0·4	0·6 (not determined)		
caprylic	—	2·7	1·7	3·9	4·7	7·6 (not determined)		
capric	—	2·3	2·2	1·7	3·5	6·0 (not determined)		
lauric	—	7·1	7·7	6·5	7·7	10·1 (not determined)		
myristic	17·1	13·6	13·7	19·5	22·5	27·5 (not determined)		
palmitic	49·5	42·7	36·4	38·2	34·6	25·0 (not determined)		
stearic	24·7	17·1	12·7	9·7	2·9	2·7 (not determined)		
arachidic	7·5	2·9	2·3	1·9	3·1	1·6 (not determined)		

* Calculated (by difference) from trisaturated glycerides determined in the whole fat.

Tables 4a and 4b show the percentage compositions (mol. acids per 100 mol.) of each group of fats separated by the preliminary crystallizations. In dealing with component glycerides it is convenient, indeed almost essential, for purposes of calculation to consider the molar rather than the weight percentages of the various constituent acyl groups. Moreover, in bovine milk fats as in other fats where acids of widely different molecular size enter into combination with glycerol, a much clearer picture of the relative number of acyl groups present as mixed glycerides in one or other combination is obtained by considering the data on a molar rather than on a weight basis. The presence in milk fats of very minor proportions of four short-chain saturated acids in addition to butyric acid renders the interpretation of the results extremely complicated and difficult, and it is therefore necessary to consider all these short-chain acids together in one group ('C_4 to C_{12}') (in the earlier study of cow milk-fat glycerides (Hilditch & Paul 1940) myristic groups were also included in this group ('C_4 to C_{14}' acids), but since myristic acid is present in some quantity not only in the milk fats but in other (e.g. depot) fats of the mammals, and in view of the somewhat greater precision of the present study, it has been thought better to leave it as an individual component).

Tables 5 *a* and 5 *b* therefore show the *increments* of various combined acid groups and of the major individual acids in each of the glyceride groups *A* to *F* of the buffalo milk fat (table 5*a*) and *A* to *H* of the cow milk fat (table 5*b*), together with the corresponding increments of the component acids of the trisaturated glycerides present in each fraction. It should be noted that, as already stated, the trisaturated glyceride contents in each of the groups *A* and *B* are calculated from the component acid data and not from the determined values. Further, the amounts of trisaturated glycerides in groups *G* and *H* of the cow milk fat were unfortunately too small to permit of separate determination of their component acids, and it was necessary to make the assumption that these were approximately as in the preceding group *F*.

TABLE 5. INCREMENTS (% MOL.) OF COMPONENT ACIDS IN THE GLYCERIDE
GROUPS AND THEIR TRISATURATED GLYCERIDES

(*a*) Indian buffalo milk fat

	A	*B*	*C*	*D*	*E*	*F*	total
% (mol.) of whole fat ...	8·5	18·4	18·6	13·7	19·5	21·3	100·0
component acid groups:							
C_4 to C_{12}	0·1	3·6	3·5	2·8	4·0	7·2	21·2
myristic	1·2	2·0	2·2	1·8	2·0	1·9	11·1
palmitic	4·0	6·3	6·5	3·3	4·1	3·3	27·5
stearic (+arachidic)	2·8	4·0	2·4	1·5	1·3	0·5	12·5
monoethenoid C_{10} to C_{16}	0·1	0·2	0·2	0·4	0·9	0·9	2·7
oleic	0·3	2·3	3·4	3·3	6·6	7·0	22·9
polyethenoid C_{18}, C_{20}, C_{22}	—	—	0·4	0·6	0·6	0·5	2·1
trisaturated glycerides:							
increments (% mol.)	7·2	10·8	8·0	6·1	4·6	3·0	39·7
component acids:							
C_4 to C_{12}	0·1	3·1	2·7	2·0	2·1	1·8	11·8
myristic	1·2	1·2	1·0	1·3	0·7	0·3	5·7
palmitic	3·5	4·2	3·1	1·9	1·3	0·6	14·6
stearic (+arachidic)	2·4	2·3	1·2	0·9	0·5	0·3	7·6

(*b*) Indian cow milk fat

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	total
% (mol.) of whole fat ...	8·8	12·7	16·3	10·1	20·2	17·7	6·1	8·1	100·0
component acid groups:									
C_4 to C_{12}	0·2	2·5	3·4	2·0	4·1	4·0	2·1	3·1	21·4
myristic	1·4	1·1	1·5	1·3	2·2	1·8	0·5	0·5	10·3
palmitic	4·0	5·3	5·6	2·5	4·7	3·1	0·7	0·6	26·5
stearic (+arachidic)	2·7	2·1	2·4	1·0	1·3	1·0	0·1	0·1	10·7
monoethenoid C_{10} to C_{16}	0·1	0·4	0·8	0·3	1·6	0·9	0·4	0·6	5·1
oleic	0·4	1·3	2·4	2·9	5·9	6·3	2·0	2·7	23·9
polyethenoid C_{18}, C_{20}, C_{22}	—	—	0·2	0·1	0·4	0·6	0·3	0·5	2·1
trisaturated glycerides:									
increments % (mol.)	7·1	7·6	7·8	2·5	4·6	2·4	1·4	1·8	35·2
component acids:									
C_4 to C_{12}	0·1	1·8	2·7	0·8	1·7	1·0	0·6	0·8	9·5
myristic	1·2	1·0	1·1	0·5	1·0	0·7	0·4	0·5	6·4
palmitic	3·5	3·3	2·8	0·9	1·6	0·6	0·3	0·5	13·5
stearic (+arachidic)	2·3	1·5	1·2	0·3	0·3	0·1	0·1	—	5·8

Comparison of the direct analysis of the whole fats with the figures obtained by summation of the data for the individual groups of glycerides shown in table 5 affords a check upon the general degree of accuracy of the detailed component acid analyses (table 6). For the most part the accordance is reasonably satisfactory, having regard to the long sequence of operations involved in the analyses of the glyceride groups separated by the low-temperature crystallization procedures.

TABLE 6. COMPARISON OF COMPONENT ACID DATA OBTAINED BY ANALYSES OF THE WHOLE MILK FATS AND OF THEIR SEPARATED GLYCERIDE GROUPS

	buffalo milk fat		cow milk fat	
	whole fat analysis % (mol.)	summation of glyceride group analyses % (mol.)	whole fat analysis % (mol.)	summation of glyceride group analyses % (mol.)
total component acids				
butyric	12·6	12·2	10·6	9·8
caproic	3·3	3·9	3·7	4·4
caprylic	0·8	0·8	1·6	2·0
capric	1·7	1·8	2·6	2·4
lauric	3·3	2·5	2·6	2·8
myristic	11·7	11·1	13·1	10·3
palmitic	26·3	27·5	28·4	26·5
stearic	11·5	11·1	6·8	9·6
arachidic	1·1	1·4	0·7	1·1
decenoic	trace	trace	0·2	0·1
dodecenoic	0·2	0·1	0·3	0·3
tetradecenoic	0·8	0·8	1·2	1·3
hexadecenoic	3·5	1·7	1·5	3·4
oleic	21·0	22·9	23·1	23·9
octadecadienoic	1·3	0·2	3·1	1·1
unsaturated C_{20} to C_{22}	0·9	2·0	0·5	1·0

Computation of the proportions of various categories of mixed glycerides present in the separated glyceride groups

If, in the first place, we consider the relative proportions of the saturated and the unsaturated acids present in any mixture of glycerides, it is clear that there are four types of mixed triglycerides possible, namely, trisaturated, disaturated mono-unsaturated, monosaturated di-unsaturated, and tri-unsaturated. In the entire milk fats, all of these types may be present; after sufficient separation by crystallization two or at most three types will be present in any one of the glyceride groups so separated. Actually, in the majority of fats (in which saturated acids of smaller molecular size than palmitic or myristic are not present) the crystallization procedure causes all the trisaturated glycerides to be assembled in the most sparingly soluble goups (e.g. *A*, *B*). Milk fats are exceptional in this respect because mixed trisaturated glycerides containing two short-chain acyl groups in addition to one of the more usual long-chain groups (e.g. dibutyromyristin or dibutyropalmitin) are as freely soluble in acetone or ether as di-unsaturated triglycerides such as palmitodiolein, and in consequence trisaturated glycerides appear throughout all

the separated glyceride groups of the milk fats. These can, however, be determined directly, and it remains to consider whether the remaining glycerides contain one, two or all of the other possible types—disaturated mono-unsaturated, monosaturated di-unsaturated, and tri-unsaturated. From the known proportions of saturated and unsaturated acids present in any one group as mixed saturated-unsaturated glycerides, the proportions of binary (but not ternary) mixtures of, for instance, disaturated mono-unsaturated and monosaturated di-unsaturated glycerides, or disaturated mono-unsaturated and tri-unsaturated glycerides can be directly calculated. Accordingly, for our purpose it is necessary to examine how far it is permissible to presuppose the substantial absence of one of these categories (e.g., tri-unsaturated glycerides). The latter are known to be the most soluble type of mixed glycerides and therefore, if present, they become concentrated in the groups of glycerides left in solution at the lowest crystallization temperatures employed. In the intermediate groups (*C*, *D*, etc.) it is legitimate to presume from these considerations of relative solubility that no tri-unsaturated glycerides are present, and the proportions of mono- and of di-unsaturated glycerides can be obtained by direct calculation as described. It will be seen from table 7 (below) that in these intermediate fractions the ratio of di-unsaturated to mono-unsaturated glycerides increases steadily with increasing solubility; the mono-unsaturated glycerides are concentrated in the earlier groups (notably *B*, *C*).

The presence of any appreciable quantities of tri-unsaturated glycerides in the most soluble groups (*F* or *H* respectively) in the present series of analyses appears very unlikely for the following reason: Calculation of the mixed saturated-unsaturated glycerides from their known contents of saturated and unsaturated acids to a mixture of mono-unsaturated and tri-unsaturated glycerides leads to increments of 14·9 % mono- and 3·4 % tri-unsaturated glycerides in group *F* of the buffalo milk fat and to increments of 3·7 % mono- and 2·6 % tri-unsaturated glycerides in group *H* of the cow milk fat. These values of about 3 % of tri-unsaturated glycerides are the maximum possible and, since it is obvious that di-unsaturated glycerides will be present in at least comparable proportions to the mono-unsaturated glycerides in these groups, it follows that the amount of tri-unsaturated glycerides actually present must be extremely small, and that it is reasonable to presume their substantial absence. In view of these considerations the relative distribution of saturated and unsaturated acyl radicals in the glyceride groups of the two milk fats may be derived from the component acid data with the results given in table 7*a*.

The proportion of the following acids (or acid subgroups) in all the glyceride fractions of both milk fats is sufficiently low to justify the conclusion that in any one triglyceride molecule only one group of the acid in question will be present: myristic, stearic (+arachidic), mono-ethenoid acids of the C_{10} to C_{16} series, and polyethenoid acids of the C_{18}, C_{20} and C_{22} series. In certain glyceride fractions, on the other hand, the proportions of either palmitic or oleic acid are such that dipalmito- or di-oleo-glycerides must be present in addition to monopalmito- or mono-oleo-glycerides. In the same way, although neither butyric nor the other short-chain acyl groups are present in any glyceride fraction to an extent which

suggests the presence of more than one of the individual acyl groups in the same triglyceride molecule, the short-chain acids (C_4 to C_{12}) taken as a whole may occur twice in some of the triglyceride molecules of certain glyceride fractions. The proportions of triglycerides containing respectively one and two groups of palmitic, oleic or the C_4 to C_{12} acids have been calculated (table 7b, c, d) from the component acid data as in the corresponding cases of the allocation of saturated and unsaturated groups as a whole (table 7a).

TABLE 7. PROPORTIONS OF VARIOUS GLYCERIDE GROUPS IN THE SEPARATED GLYCERIDE FRACTIONS OF THE MILK FATS

(a) Indian buffalo milk fat

	A	B	C	D	E	F	total
% (mol.) of whole fat ...	8·5	18·4	18·6	13·7	19·5	21·3	100·0
(a) trisaturated	7·2	10·8	8·0	6·1	4·6	3·0	39·7
disatd-mono-unsatd	1·3	7·6	9·1	2·3	5·7	11·5	37·5
monosatd-di-unsatd	—	—	1·5	5·3	9·2	6·8	22·8
tri-unsaturated	—	—	—	—	—	—	—
(b) di (C_4 to C_{12})-mono-*others*	—	—	0·2	0·1	1·6	2·5	4·4
mono (C_4 to C_{12})-di-*others*	0·2	10·8	10·2	8·2	8·7	16·7	54·8
no (C_4 to C_{12})-groups	8·3	7·6	8·2	5·4	9·2	2·1	40·8
(c) dipalmito-mono-*others*	3·5	1·7	1·2	—	—	—	6·4
monopalmito-di-*others*	5·0	15·6	17·0	10·0	12·2	9·8	69·6
no palmito-groups	—	1·1	0·4	3·7	7·3	11·5	24·0
(d) di-oleo-mono-*others*	—	—	—	2·4	4·8	2·5	9·7
mono-oleo-di-*others*	0·9	7·0	10·2	5·2	10·1	15·8	49·2
no oleo-groups	7·6	11·4	8·4	6·1	4·6	3·0	41·1

(b) Indian cow milk fat

	A	B	C	D	E	F	G	H	total
% (mol.) of whole fat ...	8·8	12·7	16·3	10·1	20·2	17·7	6·1	8·1	100·0
(a) trisaturated	7·1	7·7	7·8	2·4	4·5	2·4	1·4	1·8	35·1
disatd-mono-unsatd	1·7	5·0	6·9	5·4	7·8	7·1	1·2	1·1	36·2
monosatd-di-unsatd	—	—	1·6	2·3	7·9	8·2	3·5	5·2	28·7
tri-unsaturated	—	—	—	—	—	—	—	—	—
(b) di (C_4 to C_{12})-mono-*others*	—	—	0·4	—	0·5	0·7	0·9	1·2	3·7
mono (C_4 to C_{12})-di-*others*	0·7	7·6	9·5	5·9	11·1	10·5	4·6	6·9	56·8
no (C_4 to C_{12})-groups	8·1	5·1	6·4	4·2	8·6	6·5	0·6	—	39·5
(c) dipalmito-mono-*others*	3·4	3·2	0·7	0·4	0·2	—	—	—	7·9
monopalmito-di-*others*	5·1	9·5	15·5	6·6	13·9	9·2	2·0	1·6	63·4
no palmito-groups	0·3	—	0·1	3·1	6·1	8·5	4·1	6·5	28·7
(d) di-oleo-mono-*others*	—	—	—	1·1	2·0	3·7	1·3	1·9	10·0
mono-oleo-di-*others*	1·3	3·9	7·2	6·5	13·6	11·6	3·4	4·4	51·9
no oleo-groups	7·5	8·8	9·1	2·5	4·6	2·4	1·4	1·8	38·1

It is possible that this mode of calculation, based on the assumption that only binary mixtures (e.g. of monopalmito- with dipalmito-glycerides) are present, overlooks the presence in a few instances of very small amounts of simple triglycerides such as tripalmitin or triolein. The only case in which this is at all probable, however, is in the trisaturated glycerides of group A of the glyceride fractions, where the presence of traces of tripalmitin or tristearin cannot be excluded. There is no

evidence in our present data of the presence of the simple triglycerides tributyrin or triolein in the milk fats studied, and in this respect the findings of earlier investigators (Amberger 1913, 1918; Arup 1928) are confirmed.

Derivation of the probable constituent glycerides in each of the glyceride
groups A to H from the data in tables 5 and 7

The individual mixed glycerides present in each of the glyceride groups separated in the original crystallizations of the milk fats must now be derived as far as possible from the proportions of the component acids in each fraction (table 5) in such a manner as to conform with the various categories enumerated in table 7. In some fractions (notably the least soluble groups *A* and the most soluble groups) the number of possible components is relatively simple and the calculation is nearly unequivocal. In group *A* of the buffalo milk-fat glycerides, for example, the data correspond with this 8·5 % (mol.) of the whole fat consisting of 3·3 dipalmitostearin, 3·7 of myristopalmitostearin, 0·9 of oleopalmitostearin, 0·4 of hexadecenopalmitostearin, and 0·1 each of (C_4 to C_{12}) dipalmitin and (C_4 to C_{12})-oleo-palmitin and -stearin. Similarly, the 21·3 % of the buffalo milk fat in group *F* can be apportioned between di (C_4 to C_{12})-myristin, -palmitin, and -stearin, oleo-(C_4 to C_{12})-myristin or -palmitin, myristo- and palmito-diolein, and minor amounts of other mixed glycerides almost unequivocally. In several of the intermediate groups, although the total amounts of di- and mono-oleo-, di- and mono-palmito-, trisaturated, etc., glycerides can be derived unequivocally, the (C_4 to C_{12}), myristo-, stearo-, mono-ethenoid C_{10} to C_{16}, and polyethenoid C_{18} to C_{22} groups can be apportioned in a number of ways, all of which satisfy the conditions in table 7. In these cases the acids or acid subgroups in question have been distributed in proportion to their amounts in the various categories of glycerides consonant with the data in table 7, giving a *possible* detailed glyceride constitution for each glyceride group which, although it is considered to be the most probable, is not the only combination which satisfies the analytical data.

Each group of glycerides having thus been split up into increments of individual glycerides which conform with the analytical data in tables 5 and 7, and which (when not unequivocal) appear to be the most probable combination, summation of the respective increments leads to a statement of the possible constituent mixed glycerides in the original milk fat. The final results for both milk fats (without inclusion of the separate increments calculated for each of the glyceride groups into which the fats had been resolved) are given in table 8. No great precision can be claimed for the individual proportions suggested for some of the possible mixed glycerides. It is, however, believed that the results show the relative order of occurrence of the chief types which are present and that they form a more complete description of the glycerides of a milk fat than it has been possible to give as the result of any previous work. On the other hand, the integration of the individual glycerides suggested in table 8 into the different classes (di-oleo-, oleopalmito-, dipalmito-, etc., glycerides) shown in the subsequent table 9 (p. 203) provides data which are probably accurate to within one or two units per cent, and which it is

believed can contribute usefully to some further discussion of the possible origin of the milk fats.

TABLE 8. POSSIBLE COMPONENT GLYCERIDES OF THE MILK FATS

	buffalo % (mol.)	cow % (mol.)
trisaturated:		
di (C_4 to C_{12})-myristin	0·7	0·8
di (C_4 to C_{12})-palmitin	3·1	1·7
di (C_4 to C_{12})-stearin	0·6	—
(C_4 to C_{12})-myristo-palmitin	10·7	12·5
(C_4 to C_{12})-dipalmitin	2·2	2·4
(C_4 to C_{12})-myristo-stearin	1·2	0·7
(C_4 to C_{12})-palmito-stearin	12·7	7·8
myristodipalmitin	0·3	0·5
myristopalmitostearin	4·3	4·8
dipalmitostearin	3·9	3·9
disaturated-mono-unsaturated:		
di (C_4 to C_{12})-olein	—	1·1
(C_4 to C_{12})-myristo-olein	5·8	4·6
(C_4 to C_{12})-palmito-olein	9·3	5·8
(C_4 to C_{12})-stearo-olein	2·1	3·6
myristopalmito-olein	6·4	6·3
dipalmito-olein	—	1·1
myristostearo-olein	2·5	1·0
palmitostearo-olein	10·0	9·7
(C_4 to C_{12})-palmito-monoethenoid*	0·6	1·9
monoethenoid-palmito-stearin	0·4	0·4
(C_4 to C_{12})-palmito-polyethenoid*	0·4	0·7
monosaturated-di-unsaturated:		
(C_4 to C_{12})-diolein	4·1	6·4
myristo-diolein	1·4	—
palmitodiolein	4·3	3·6
(C_4 to C_{12})-monoethenoid*-olein	2·9	6·5
palmito-monoethenoid*-olein	4·2	6·5
(C_4 to C_{12})-polyethenoid*-olein	2·8	4·0
palmito-polyethenoid*-olein	3·1	1·7

* monoethenoid (C_{10} to C_{16}) acids; polyethenoid (C_{18}, C_{20}, C_{22}) acids.

DISCUSSION

The figures in table 8, irrespective of the precise degree of their numerical accuracy in any given instance, emphasize the extremely large number of individual mixed triglycerides which go to make up a bovine (or, indeed, any ruminant) milk fat. The most abundant glyceride forms can be reduced to the seven classes shown in table 9a, whilst considerable insight into the general pattern of milk-fat glycerides can be gained by examination of the relative occurrence therein of palmitic, oleic, and the short-chain acyl groups. Table 9 illustrates these features from the two milk fats primarily dealt with in this communication, together with corresponding data for an English cow milk fat taken from the less complete analysis made by Hilditch & Paul (1940). In comparing these three fats, it should be remembered that the proportion of oleic acid and of unsaturated acids as a whole in the total fatty acids differs in each case, the molar percentages of oleic and of total un-

saturated acids being respectively English cow 31, 41, Indian cow 24, 31, Indian buffalo 23, 28; consequently the proportions of mono- and of di-oleo-glycerides vary accordingly in the three fats.

TABLE 9. CONSTITUENT GLYCERIDES OF BUFFALO AND COW MILK FATS

		Indian buffalo (present work) % (mol.)	Indian cow (present work) % (mol.)	English cow (Hilditch & Paul 1940) % (mol.)
(a) Some of the more abundant constituents				
triglyceride types:				
palmitic lower acyl	lower acyl	14	14	6–7
palmitic lower acyl	stearic	17	13	9
palmitic oleic	lower acyl	16	12	30–22
palmitic oleic	stearic	10	10	8–17
palmitic oleic	oleic	4	4	17–4
oleic stearic	lower acyl	5	5	12–6
oleic ⸱ oleic	lower acyl	6	6	0–10
(b) Distribution of palmitic groups				
glycerides with 2 palmitic groups		6	8	5
glycerides with 1 palmitic group		70	64	65
glycerides with no palmitic group		24	28	30
(c) Distribution of oleic groups				
glycerides with 2 oleic groups		10	10	17
glycerides with 1 oleic group		49	52	64
glycerides with no oleic group		41	38	19
(d) Distribution of lower saturated acyl groups (C_4 to C_{14})				
glycerides with 2 lower acyl groups		22	21	11–17
glycerides with 1 lower acyl group		52	52	52–46
glycerides with no lower acyl group		26	27	37

It is evident that in general pattern the glycerides of all three milk fats have much in common, and this is especially well marked in regard to the distribution of palmitic groups in the glyceride molecules. The circumstance that there is an equally well-marked parallelism between the proportions and distribution of palmitic acid in the milk and depot (body) fats of a range of mammals is still more interesting. Data are still very incomplete in this field, but table 10 offers a comparison of the palmitic acid contents of human milk and body fats (Hilditch & Meara 1944 *a, b*; Cramer & Brown 1943) with those of the pig (de la Mare & Shorland 1944; Hilditch, Lea & Pedelty 1939), sheep and goat (Hilditch & Jasperson 1944; Hilditch & Pedelty 1941; Dhingra & Sharma 1938) and cow and buffalo (Hilditch & Longenecker 1937; Hilditch & Paul 1938; Hilditch & Murti 1940; Achaya & Banerjee 1946). In those cases in which the glycerides have also been studied the proportions of mono- and of di-palmito-glycerides are included.

The palmitic acid content of the milk fat of any of the animals listed in table 10 is usually very similar to, but a few units per cent less than, that of the depot fat of the same animal; and, so far as data are available, it appears that its distribution

either as monopalmito- or as dipalmito-glycerides in both milk and body fats is on the whole broadly similar. In this tendency towards constant proportions palmitic acid stands in marked contrast to all the other saturated acids present in milk fats. Thus it will be seen from table 11 (p. 206) that, whereas sow milk fat contains only traces of acids of smaller molecular size than palmitic, human milk fat has appreciable proportions of myristic and lauric and a little capric acid; whilst the milk fats of ruminants contain the whole range of even-numbered saturated acids down to butyric. The clearly defined and relatively constant palmitic acid pattern of both milk and depot glycerides of the mammals appears to lend additional support to the view (Banks & Hilditch 1931, 1932; Hilditch & Sleightholme 1930, 1931; Hilditch 1947, pp. 303–306) that both classes of fats may

TABLE 10. DISTRIBUTION OF PALMITIC ACID IN ANIMAL MILK
AND DEPOT FATS

| | % (mol.) of total acids | | distribution in glycerides | | | |
| | | | monopalmito- | | dipalmito- | |
animal	milk fat	depot fat	milk fat	depot fat	milk fat	depot fat
pig	28	30	—	75–80	—	7–13
human	23	28	60	—	6	—
sheep	22	27–30	—	75	—	9–16
goat	24	27	—	—	—	—
cow (English)	23–24	27–30	65	70	5	13
cow (Indian)	28	31–33	64	—	8	—
buffalo (Indian)	27	31	70	—	6	—

be produced in the animal from precursor glycerides which are primarily of a mainly unsaturated nature and of which palmitodioleins or some similar palmito-di-unsaturated glycerides form the predominating component. This hypothesis, so far as milk fats are concerned, is in accordance with experimental findings (Graham, Jones & Kay 1936; Maynard & McCay 1938; Shaw & Petersen, 1938, 1940) that, during passage through the mammary gland of the cow, glycerides (but not other lipids) present in the blood disappear in amounts sufficient to account for the milk fat secreted; but other views, which postulate a non-fatty precursor for the milk fat, may be considered before we return to further discussion of this hypothesis in the light of the present and of earlier work.

To attempt to deduce from the chemical constitution of a metabolic end-product the probable course of its synthesis in the animal is in any case difficult, since it is rarely possible to obtain direct knowledge of intermediate stages in the process, and the conclusions depend almost wholly upon evidence of a circumstantial nature. On the other hand, the alternative hypotheses which have been put forward for the biosynthesis of milk fat from a carbohydrate or other non-fatty precursor, whilst based upon certain other experimental findings, have never taken into account the specific constitution of the final product. This is unfortunate, since no hypothesis of milk-fat formation can be acceptable if it does not lead to specific mixtures of milk fatty acids or glycerides which are in conformity with the abundant data pub-

lished on this subject during the past twenty years. Judged by this criterion, the alternative suggestions which have been put forward do not accord with the known structure of milk fats:

(1) *Milk-fat synthesis from glucose (or other carbohydrate)*. The respiratory quotient of the lactating mammary gland has been observed to be greater than unity (e.g. Graham, Houchin, Peterson & Turner 1938), and this has been taken to indicate the probable conversion therein of carbohydrate into milk fat. Shaw & Petersen (1940), however, suggested that some other explanation must be sought for the high respiratory quotient. Leathes & Raper (1925, pp. 177, 178) emphasize that values of the respiratory quotient 'merely indicate the type of reaction which is taking place, and do not identify it necessarily as being the conversion of fat into carbohydrate, or vice versa'.

If the acids of milk fat were, like those (for example) of seed fats, synthesized directly from carbohydrate, it would be expected that their distribution in the glycerides would be on the same pattern as that which has been found to obtain in all but two classes of natural fats, vegetable or animal. This distribution can be readily illustrated by the statement that an acid which forms about one-third or more of the total fatty acids will be found to be present in practically all the triglyceride molecules. As a corollary it follows that, in fats with high proportions of saturated acids, the proportion of trisaturated glycerides in a fat is confined to that which remains after all the unsaturated acids have been included in the form of mono-unsaturated glycerides. The only two classes of natural fats which do not follow this generalization are animal body fats which contain high proportions of stearic acid, and animal milk fats which contain large amounts of the short-chain fatty acids. Thus, the cow and buffalo milk fats described in this paper would, if their glyceride structure followed the general rule, contain respectively not much more than 7 and 17 % of trisaturated glycerides, whereas the observed figures are 35 and 40 %. If, as in other natural fats, the milk fatty acids had all been synthesized from carbohydrate and then esterified into glycerides, one would be forced to conclude, with Norris & Mattil (1946), that there is a fundamental difference in the esterifying action of most of the lipolytic enzymes of plants and animals on the one hand, and of those of a few specific animal tissues on the other. This unlikely supposition is unnecessary if the typical glyceride structure of animal depot and milk fats is the result of further transformation of oleo-glyceride precursors which have already been synthesized on the normal 'evenly-distributed' pattern at an earlier stage.

(2) *Milk-fat synthesis from acetic acid produced in the rumen*. Since the production of short-chain acids is far more marked in the milk fats of ruminants than elsewhere, the suggestion (Barcroft, McAnally & Phillipson 1944) that this may be closely connected with the presence in the blood stream of acetic and other lower fatty acids produced by anaerobic bacterial action upon carbohydrates in the rumen appears more attractive than that of their specialized synthesis from carbohydrate in the mammary gland itself. Folley & French (1949) have, indeed, offered evidence to show that in ruminants the short-chain components of the milk fat may be synthesized from acetate rather than from carbohydrate. Again, however, this view

(which rests solely on the respiratory quotient) is not in harmony with the typical structures of milk fats as given above; and it is unfortunate that no information has been given as to the composition of the fats produced experimentally, although admittedly the small amounts available would necessitate only an abridged form of component acid investigation. Consideration of the different proportions of short-chain acids produced in the milk fats of different animals (table 11) emphasizes their specific predominance in ruminant milk fat, but appears to us not to support the view that their direct synthesis from acetate has taken place to any marked extent.

TABLE 11. PROPORTIONS (% MOL.) OF SHORT-CHAIN SATURATED ACIDS
IN DIFFERENT MILK FATS

milk fat	satuarated								unsaturated	
	C_4	C_6	C_8	C_{10}	C_{12}	C_{14}	C_{16}	C_{18}	oleic	other
pig	←————2————→					2	28·5	6	35	26·5
human	—	—	—	2	8	9	23	9	34	15
horse	1	2	4	8	7	7	15	3	37	16
camel	6	2	1	2	6	8	28	10	34	3
sheep	8	5	3	6	5	10	22	11	22	8
goat	7	5	4	13	7	12	24	5	17	6
cow (English)	10	4	2	3	2	8	24	11	29	7
cow (Indian)	11	4	2	3	2	13	28	7	23	7
buffalo (Indian)	12	4	1	2	3	11	27	12	21	7

	total C_4 to C_{12} fatty acids	chief short-chain component acids
pig	2	—
human	10	lauric
horse	22	capric, lauric
camel	17	butyric, lauric
sheep	27	butyric, capric
goat	36	butyric, capric
cow (English)	21	butyric
cow (Indian)	22	butyric
buffalo (Indian)	22	butyric

Whilst detailed component acid analyses have been given only for the milk fats of the animals mentioned in table 11, general analytical characteristics (Reichert and Polenske values) show clearly that short-chain saturated acids are absent from the milk fats of mice, dogs and cats (as well as pigs), whilst those of asses and rabbits are similar to camel milk fats in Reichert values. In many animals, therefore (as in the pig), the milk fats are evidently not very different from their other body fats; at the other extreme, the ruminant milk fats contain 20 to 30 % of short-chain acids, in which butyric acid is the chief component. Intermediate between these groups fall certain other milk fats, notably human and mare milk fat, in which the most abundant short-chain acids are those with 10 and 12 carbon atoms. If these acids are indeed synthesized in the lactating gland from either carbohydrate or acetate, it follows that in these two instances of non-ruminants, the short-chain acids so synthesized (although perhaps less in total amount than that usually produced by ruminants) consist predominantly of acids with 12 or 10 carbon atoms —in other words, the synthetic process has been much more intensive in that

coupling of five or six C_2 units has been effected, whereas in the ruminants most of the short-chain acids are derived only from two or three C_2 units (butyric and caproic acids).

This relationship is inconsistent with the hypothesis that synthesis from a C_2 (acetate) unit is specially favoured in the lactating glands of ruminants. If, however, the data in table 11 are considered in the light of possible transformation of oleic (and perhaps other unsaturated) acyl groups of glycerides entering the mammary gland in the blood, a much more consistent picture results; in many animals little change may occur, in others a certain degree of chain-shortening of oleo-glycerides is apparent but only proceeds to a relatively small extent and (as in human milk fat) leads only to increase in the myristic (C_{14}) content and to production of some C_{12} and less C_{10} acids. In the milk fats of the ruminants, however, such changes proceed to a much more complete degree, the quantity of the short-chain acids is much greater, and the chain shortening extends as far as production of butyric glycerides in greater amount than any of the other short-chain derivatives. This may well be due to the specific acetic and lower fatty acid content of the rumen upon which Folley *et al.* have focused attention, but we would suggest that the effect lies in some consequent alteration (e.g. in pH value) in the condition of the mammary gland which favours the action of the enzymes concerned in a chain-shortening mechanism, rather than in direct synthesis of lower fatty glycerides from acetate or other non-fatty precursor.

It will again be noticed in table 11 that, except in one instance (mare milk fat), the palmitic acid content of all the milk fats is but little reduced from that of a (sow) milk fat in which no short-chain acids are present. This circumstance would allow the introduction of, at most, only small additional proportions of the separate synthetic short-chain glycerides from acetate suggested by Folley *et al.*

We return, in conclusion, to the hypothesis that the short-chain acyl glycerides of milk fat result almost wholly from oxidation and reduction processes which operate mainly upon the unsaturated (chiefly oleic) groups originally present in the blood glycerides during their passage through the lactating gland. The evidence, admittedly largely indirect, in favour of this theory must, we feel, now be regarded as considerable, and is recapitulated briefly below:

(1) In all milk fats which contain lower saturated fatty acids, the molar proportions of the latter bear an inverse relationship to that of the oleic acid, the sum of these tending towards a constant figure, whilst the remaining major component acid—palmitic—also approaches constancy throughout.

(2) In these milk fats the content of trisaturated glycerides rises much more rapidly, as the proportion of unsaturated acids in the total acids declines, than in any other natural fats with the sole exception of stearic-rich animal depot fats. When the proportion of trisaturated glycerides is plotted against the percentage of saturated acids in the total acids, the observed points fall close to a curve which intersects the horizontal axis at a point corresponding to a molar content of 30 to 35 % of saturated acids in the total acids (Hilditch 1947, pp. 297–303).

(3) Precisely the same general features have been found to obtain in those animal depot fats which contain large proportions of stearic acid, trisaturated glycerides

appearing to a much larger extent than in other fats so soon as stearic acid is present in quantity in addition to the 30 % or thereabouts of palmitic acid which is an almost constant characteristic of all animal fats. In these stearic-rich fats it is suggested that reduction of precursor oleo-glycerides subsequent to their synthesis accounts for the specific glyceride structures observed.

(4) When a lactating cow is starved (Smith & Dastur, 1938), or suffers from ketosis (Shaw 1941 a, b; Shaw, Powell & Knodt, 1942), the proportion of short-chain acyl constituents falls very markedly, whilst the oleic acid content rises to a corresponding extent, the palmitic acid content remaining practically unchanged. On resuming normal feeding after fasting, or on recovery from ketosis, the composition of the milk fats from the cows rapidly reverts to its normal content of lower fatty acids and a correspondingly lowered content of oleic acid.

(5) When vegetable oils rich in linoleic and linolenic acid are included in the diet, the cow milk fat is but little affected in composition, and these polyethenoid acids do not appear to any notable extent in the milk fat. When, however, cod-liver oil (which contains considerable proportions of tetra- and penta-ethenoid acids of the C_{20} and C_{22} series) is similarly administered, these highly unsaturated acids are found to the extent of several per cent in the milk fat, and the content of lower fatty acids is reduced to half or less of the normal amount (Hilditch & Thompson 1936). This is consistent with adsorption of the highly unsaturated fish-oil glycerides at the enzyme system in the lactating gland which, however, is not able to metabolize them owing to their different chain length as compared with the usual unsaturated C_{18} chain of oleic and related acids.
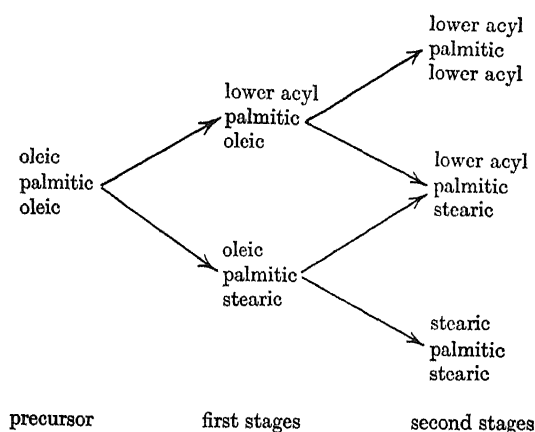
(6) The lower saturated acids of milk fats are invariably accompanied by small proportions (decreasing as the carbon chain shortens) of mono-ethenoid acids of the C_{16}, C_{14}, C_{12} and C_{10} series (but not below C_{10}), in which the ethenoid bond occupies the same position ($-CH:CH.[CH_2]_7.CO_2-$) relative to the carboxyl groups as in oleic acid (Hilditch & Longenecker 1938). The occurrence of this series of minor unsaturated acids is peculiar to milk fats in which the lower saturated acids are also prominent, and is consistent with the presence of small amounts of transformed oleo-glycerides which have escaped complete saturation to lower saturated glycerides.

(7) The more detailed information on the constituent glycerides of bovine milk fats obtained in the course of the present work permits some further evidence to be added, especially in regard to the more or less constant pattern of palmito-glycerides which persists throughout these fats and the corresponding depot fats:

(i) The proportions of mono- and di-palmito-glycerides which have been observed in bovine milk and depot fats (table 10), in relation to the content of palmitic acid (c. 24 to 30 %) in their total fatty acids, are of the order characteristic of the usual 'evenly distributed' type of a fatty acid throughout the glyceride molecules of a fat. In contrast, the distribution of the lower acyl groups and oleic groups in the ruminant milk-fat glycerides approaches more closely to values calculable by the theory of probability than to those for 'evenly' distributed types. This is what would be expected, since attack upon oleo- (and perhaps other un-

saturated) groups in the mixed glycerides by an oxidation-reduction mechanism would be random, rather than selective, in character.

(ii) The conversion of a palmitodiolein precursor into a trisaturated glyceride by conversion of the oleo-groups either into stearo-groups by hydrogenation, or into lower acyl groups by chain-shortening would obviously involve two stages which may be indicated diagrammatically as follows:

```
                                              lower acyl
                                              palmitic
                                              lower acyl
                       lower acyl
                       palmitic
   oleic               oleic                  lower acyl
   palmitic                                   palmitic
   oleic                                      stearic
                       oleic
                       palmitic
                       stearic
                                              stearic
                                              palmitic
                                              stearic

   precursor           first stages          second stages
```

The observed proportions of each of these stages in the monopalmito-glycerides of the two milk fats and the two depot fats cited above are shown in table 12, together with the percentages of the three categories of monopalmito-glycerides to which they correspond.

TABLE 12. POSSIBLE DERIVATION OF BUFFALO AND COW MILK
FAT GLYCERIDES FROM A PALMITODIOLEIN PRECURSOR

				milk fats		depot fats	
				buffalo % (mol.) (obs.)	cow % (mol.) (obs.)	cow (Indian) % (mol.) (obs.)	cow (English) % (mol.) (obs.)
palmito-glyceride types:							
oleic	palmitic	oleic	(precursor)	4	4	17	23
oleic	palmitic	stearic	(first stage)	10	10	38	32
oleic	palmitic	lower acyl	(first stage)	16	12	—	—
stearic	palmitic	lower acyl	(second stage)	17	13	—	—
lower acyl	palmitic	lower acyl	(second stage)	14	14	—	—
stearic	palmitic	stearic	(second stage)	—	—	12	6
percentage of palmitodiolein (precursor) converted to first and second stages							
unchanged precursor				6	8	25	38
first stage				43	41	57	52
second stage				51	51	18	10

It is clear that the supposed change has taken place to a much greater extent in the ruminant milk fats than in their depot fats. Apart from the evidently greater liability of these changes to occur in the case of ruminant milk-fat production, it

will be noted that in the latter there are two ways of reaching the first stage, and that further transformation of either of these first stages can also take place in two ways, leading to three different categories of palmito-disaturated glycerides in milk fats. In the postulated bio-hydrogenation process which alone would be operative in depot fats, there is only one way in which each stage can be reached.

It is therefore considered that the figures in table 12 accord with the hypothetical scheme put forward, and that they offer additional support for the view that animal milk fats and stearic-rich depot fats may be produced for the most part from pre-formed oleo-glycerides by the general mechanisms to which reference has been made.

In concluding this discussion which, as remarked at the outset, is bound to rely very largely on inferential evidence as to the mechanism of milk- and depot-fat production in the animal based upon careful study of the end-products in question, the suggestion is offered that the more direct approach (e.g. by *in vitro* studies with gland or tissue slices) should be extended to include the study not only of glucose or acetate, but also of triolein, palmitodiolein, or suitable natural fatty oils containing high proportions of these glycerides, as substrates. It would appear that so far only non-fatty precursors have been examined in this way; the indications that fatty glycerides may themselves be the precursors of such specific fats as those of milk are surely sufficient to justify physiological studies which might either verify or disprove the hypothesis.

REFERENCES

Achaya, K. T. & Banerjee, B. N. 1946 *Biochem. J.* **40**, 664.

Amberger, C. 1913 *Z. Untersuch. Nahr.- u. Genussm.* **26**, 65.

Amberger, C. 1918 *Z. Untersuch. Nahr.- u. Genussm.* **35**, 313.

Arup, P. 1928 *Analyst.* **53**, 641.

Banks, A. & Hilditch, T. P. 1931 *Biochem. J.* **25**, 1168.

Banks, A. & Hilditch, T. P. 1932 *Biochem. J.* **26**, 298.

Barcroft, J., McAnally, R. A. & Phillipson, A. T. 1944 *J. Exp. Biol.* **20**, 120.

Bjarnason, O. & Meara, M. L. 1944 *J. Soc. Chem. Ind., Lond.*, **63**, 61.

Cramer, D. L. & Brown, J. B. 1943 *J. Biol. Chem.* **151**, 427.

De la Mare, P. B. D. & Shorland, F. B. 1944 *Nature*, **153**, 380.

Dhingra, D. R. 1933 *Biochem. J.* **27**, 851.

Dhingra, D. R. 1934 *Biochem. J.* **28**, 73.

Dhingra, D. R. & Sharma, D. N. 1938 *J. Soc. Chem. Ind., Lond.*, **57**, 369.

Dunn, H. C., Hilditch, T. P. & Riley, J. P. 1948 *J. Soc. Chem. Ind., Lond.*, **67**, 199.

Folley, S. J. & French, T. H. 1949 *Nature*, **163**, 174; *Biochem. J.* **45**, 117.

Graham, W. R., Jones, T. S. G. & Kay, H. D. 1936 *Proc. Roy. Soc.* B, **120**, 330.

Graham, W. R., Houchin, O. B., Peterson, V. E. & Turner, C. W. 1938 *Amer. J. Physiol.* **122**, 150.

Haller, A. 1906 *C.R. Acad. Sci., Paris*, **143**, 657, 803.

Hilditch, T. P. 1947 *The chemical constitution of natural fats*, 2nd ed. London: Chapman and Hall, Ltd.

Hilditch, T. P., Holmberg, J. & Meara, M. L. 1947 *J. Amer. Oil. Chem. Soc.* **24**, 321.

Hilditch, T. P. & Jasperson, H. 1944 *Biochem. J.* **38**, 443.

Hilditch, T. P., Laurent, P. L. & Meara, M. L. 1947 *J. Soc. Chem. Ind., Lond.*, **66**, 19.

Hilditch, T. P. & Lea, C. H. 1927 *J. Chem. Soc.* p. 3106.

Hilditch, T. P., Lea, C. H. & Pedelty, W. H. 1939 *Biochem. J.* 33, 493.

Hilditch, T. P. & Longenecker, H. E. 1937 *Biochem. J.* 31, 1805.

Hilditch, T. P. & Longenecker, H. E. 1938 *J. Biol. Chem.* 122, 497.

Hilditch, T. P. & Maddison, L. 1940 *J. Soc. Chem. Ind., Lond.*, 59, 162.

Hilditch, T. P. & Maddison, L. 1941 *J. Soc. Chem. Ind., Lond.*, 60, 258.

Hilditch, T. P. & Maddison, L. 1942 *J. Soc. Chem. Ind., Lond.*, 61, 169.

Hilditch, T. P. & Maddison, L. 1948 *J. Soc. Chem. Ind., Lond.*, 67, 253.

Hilditch, T. P. & Meara, M. L. 1944a *Biochem. J.* 38, 29.

Hilditch, T. P. & Meara, M. L. 1944b *Biochem. J.* 38, 437.

Hilditch, T. P. & Murti, K. S. 1940 *Biochem. J.* 34, 1301.

Hilditch, T. P. & Pathak, S. P. 1947 *J. Soc. Chem. Ind., Lond.*, 66, 421.

Hilditch, T. P. & Paul, S. 1938 *Biochem. J.* 32, 1775.

Hilditch, T. P. & Paul, S. 1940 *J. Soc. Chem. Ind., Lond.*, 59, 138.

Hilditch, T. P. & Pedelty, W. H. 1941 *Biochem. J.* 35, 932.

Hilditch, T. P. & Sleightholme, J. J. 1930 *Biochem. J.* 24, 1098.

Hilditch, T. P. & Sleightholme, J. J. 1931 *Biochem. J.* 25, 507.

Hilditch, T. P. & Thompson, H. M. 1936 *Biochem. J.* 30, 677.

Jack, E. L. & Henderson, J. L. 1945 *J. Dairy Sci.* 28, 65.

Jack, E. L., Henderson, J. L. & Hinshaw, E. B. 1946 *J. Biol. Chem.* 162, 119.

Leathes, J. B. & Raper, H. S. 1925 *The fats*, 2nd ed. London: Longmans Green and Co.

Maynard, L. A., McKay, C. M., Ellis, G. H., Hodson, A. Z. & Davis, G. K. 1938 *Mem. Cornell Univ. Agric. Exp. Sta.* no. 211.

Norris, F. A. & Mattil, K. F. 1946 *Oil and Soap*, 23, 289.

Shaw, J. C. 1941a *J. Dairy Sci.* 24, 502.

Shaw, J. C. 1941b *J. Dairy Sci.* A, 3, 145.

Shaw, J. C. & Petersen, W. C. 1938 *J. Dairy Sci.* 21, 122.

Shaw, J. C. & Petersen, W. C. 1940 *J. Dairy Sci.* 23, 1045.

Shaw, J. C., Powell, R. C. & Knodt, C. B. 1942 *J. Dairy. Sci.* 25, 909.

Smith, J. A. B. & Dastur, N. N. 1938 *Biochem. J.* 32, 1868.

T

The influence of geographic isolation on the skull of the green monkey (*Cercopithecus aethiops sabaeus*)

I. A comparison between the teeth of the St Kitts and the African green monkey

By E. H. ASHTON AND S. ZUCKERMAN, F.R.S.

Department of Anatomy, University of Birmingham

A quantitative analysis has been made of the dental dimensions and indices of (*a*) a collection of skulls of the green monkey (*Cercopithecus aethiops sabaeus*) brought together from the island of St Kitts in the West Indies, and (*b*) a corresponding group of green monkey skulls collected in West Africa. The green monkey was introduced into St Kitts some 300 years ago; the African green monkey is the modern representative of its parent stock.

Comparison shows that (*a*) the teeth of the green monkey in the West Indies are bigger than those of its present-day African cousin; (*b*) that there has been a decrease in the variance of the linear dimensions of the cheek teeth; and (*c*) that there has been an increase in the occurrence of such dental abnormalities as malposition, numerical variations, and variations in the number of roots of the third molar. These changes can be explained as an effect of selection acting on the original genetic constitution of the stock of green monkeys which became established in the Caribbean.

INTRODUCTION

Primates other than tarsioids and lemuroids first appear in Lower Oligocene formations, and the Miocene contains remains that are representative of most modern forms. With the exception of man, who first appears in clearly defined form in the early Pleistocene, the Primate group thus went through its major differentiation, in common with most other mammals, in the earlier part of the Tertiary. The speed with which this differentiation occurred cannot, however, be estimated from the very small number of fossils which represent these early phases of evolution. Information on the general question of the rate of change in Primates must therefore be extracted from such observations as can be made about changes in extant forms, it being accepted at the start that evolution in Primates, as in other organisms, is a measure of the effects of selection and isolation upon a variable genetic system.

Simpson (1944), who has closely considered the problem of rates of evolution, refers to only two corresponding pieces of information for the Mammalia. The first relates to a stock of the marine seal *Phoca vitulina* which has been isolated in the Upper and Lower Seal Lakes in Canada since the end of the last glacial epoch—a period judged to be from 3000 to 8000 years. The length of each generation in this species is not established, but is estimated to be from five to ten years. These seals have therefore been isolated for a period of from 300 to 1600 generations. During this time, the lake and oceanic stocks have diverged in a number of characters which Doutt (1942) considers adequate to warrant the formation of the new subspecies *P. vitulina mellonae* to accommodate the lake form. He claims that

although it is in general more like the Pacific subspecies *P. v. richardii* than the Atlantic form *P. v. concolor*, in some anatomical details it resembles other species of this genus.

The second piece of evidence relates to groups of small rodents isolated on islands or in mines. Jameson (1898), for example, observed that the colour of a sample of mice captured on a sandy island in Dublin Bay, which was almost completely separated from the mainland, was lighter than that of the mainland species. From a study of topographical charts, he estimated that the island stock had been isolated for a maximum period of approximately 120 years or 300 generations.

There is one published piece of evidence of a corresponding kind relating to the Primates. Ashley Montagu (1937) notes that five or eight *Colobus* monkeys (*C. kirkii*) from a stock isolated on the island of Zanzibar had patent metopic sutures. He considers that this figure is significantly high, and attributes it to the results of geographical isolation. However, he also records that the metopic suture was patent in two of four specimens of the mainland variety *C. ferrugineus*. Our own analysis of these figures suggests that they are insufficient to prove that the discrepancy between the two groups of observations is due to causes other than variations in random sampling from a single homogeneous population. ($\chi^2 = 0\cdot043$ with one degree of freedom; $P = 0\cdot9$–$0\cdot8$.)

The present paper is the first of a series which will analyze the changes that have occurred in the cranial characters and teeth of a group of African green monkeys, *Cercopithecus aethiops sabaeus*, which have been isolated on the island of St Kitts in the West Indies for a period of some 300 years.

THE GREEN MONKEY OF ST KITTS

The most recent revision of the genus *Cercopithecus* is that of Schwarz (1928). The species *C. aethiops* contains the majority of the green-bodied African monkeys and is divided into sixteen subspecies, which on the east range from South Africa to Abyssinia and the Sudan, and which on the west extend from the Congo to Senegambia and the Niger. Some of the subspecies, for example, *C. aethiops sabaeus*, the green monkey; *C. ae. aethiops*, the Abyssinian monkey; and *C. ae. pygerythrus*, the vervet monkey of South Africa, are fairly close to each other in size and general appearance. Others, such as *C. ae. cynosurus*, the malbrook monkey; and *C. ae. tantalus*, the tantalus monkey, are more massive animals.

C. ae. sabaeus has a wide distribution in West Africa, extending from Senegambia, in French West Africa, to the Niger in the north, and south-eastwards into the Congo (Elliot 1912; Forbes 1894). Elliot (1912) states that a variant of this subspecies is to be found in the Congo district.

Green monkeys have for many hundreds of years been exported as pets to all parts of the world. In two or three islands of the West Indies, and according to Schlegel (1876) and Sclater (1893) in St Jago in the Cape Verde Islands, the animal has escaped and established itself in its new surroundings. The only one of these accidentally established wild groups of green monkeys that is at all well known lives on the island of St Kitts in the British West Indies.

St Kitts, or St Christopher, is one of the Leeward Islands. It was discovered by Columbus in 1493, but was not settled until 1623, when an English party under Sir Thomas Warner made their home there. French colonists reached the island about the same time, and in 1628 it was divided between England and France. In 1713 the whole island was ceded to the British Crown as one of the terms of the Treaty of Utrecht (Campbell 1753).

Slaves from West Africa were introduced by the earliest settlers to help to run the West Indian sugar plantations. The export of pet monkeys was simply a by-product of the slave trade, and it was in this way that the green monkey was introduced to St Kitts some time during the seventeenth century. Not long after, during a period of quarrel between the British and French settlers, monkeys escaped from abandoned French houses and began to live wild. This observation is recorded by Père Labat (1722), who visited St Kitts in 1700. He does not, however, give any indication of the origin of the monkeys, which he describes as living wild at the time of his visit, and as being a considerable menace to the plantations. They had, in fact, become such a pest that in 1682, under the Governorship of Sir Richard Dutton, they were officially declared vermin, and a bounty was offered for every one that was killed.

Labat does not specifically state that the animal he saw was the green monkey of West Africa. Except in one detail, however, his description suggests that it was undoubtedly a catarrhine, and not a platyrrhine species. He writes of its depredations of plantations in precisely the same terms as one encounters in descriptions of African monkeys. He indicates, too, that the monkey is interested not only in tree fruits, but also in potatoes and other ground crops. His description also indicates that the animals he saw had cheek pouches, which is diagnostic of a catarrhine as opposed to a platyrrhine species. The only point in which his description does not quite fit the picture of a catarrhine monkey is in the statement that one of the four monkeys which were shot in a hunt in which he took part had a young one on its back. The young of catarrhine monkeys generally, although not invariably, cling to the chest or belly of their mothers (Zuckerman 1933).

Labat's description of the monkey is the most informative which appeared until Sclater in 1866 recorded the fact that specimens which he diagnosed as the West African green monkey had been sent from St Kitts to the London Zoological Gardens.

While the direct evidence does not therefore suggest that the green monkey has lived for more than 100 years on St Kitts, the indirect provides a fairly clear indication that the period may be as long as 300 or, taking a generation as from three to four years, as many as from 75 to 100 generations.

MATERIAL AND METHODS

Skulls

The ninety-five St Kitts monkeys' skulls used in the present study form a collection that has been brought together at the Royal College of Surgeons, London,

by Sir Frank Colyer. They have been compared with forty-nine skulls of the African green monkey, *C. ae. sabaeus*, belonging to the British Museum of Natural History, London, and the Royal College of Surgeons.

Diagnosis of sex

It is widely believed that the teeth of the male monkey are generally bigger than the female, and it was therefore considered necessary to separate the sexes in the study of the green monkey, in spite of the fact that in this species sexual dimorphism is in general not very pronounced. In the baboon, where it is very marked, many dental dimensions were found to be bigger in the male than in the female in a statistical survey of skulls of the species *Papio porcarius*, whose sexes were known from field records (Ashton & Zuckerman, unpublished observations).

The sexes of very few of the skulls of either of the two groups of green monkey studied were, however, known from field records. Those of most of the St Kitts group had been previously determined by inspection by Sir Frank Colyer, and in all except one case, the sexes recorded in the catalogue were accepted as stated. Those of the control African group of skulls were determined in the usual way by a qualitative examination of:

(a) Size.

(b) The degree of development of the muscular markings and crests.

(c) The size of the canine teeth (cf. Zuckerman 1926, 1928).

Comparison of the means of those dimensions not affected by attrition, and of the variances of these dimensions in the teeth of the monkeys sexed in this way shows that in all the dimensions examined except one, every male tooth is larger than the corresponding female tooth, the differences being in many cases statistically significant. The fact that this observation agrees so closely with the subdivision of the skulls into male and female groups on the basis of the usual cranial criteria provides a reasonable check on the adequacy of our method of sex diagnosis.

Skulls younger than the dental age represented by the eruption of the permanent incisors could not be sexed, and all deciduous teeth were therefore included in one group.

Measurement

The measurements taken in this study, and the techniques of measurement, are the same as those described in an analysis of the dimensions and indices of anthropoid teeth (Ashton & Zuckerman 1950). Indices were, however, computed only for the molar teeth of the green monkey.

Diagnosis of age

It is clear that some of these measurements, chiefly those of the height of the incisors and canines, are affected by attrition and therefore indirectly by age. In comparing the teeth of the island and mainland stocks it was therefore considered useful in the first place to compare groups of animals of similar age. This was done for the permanent teeth, but because of the small number of specimens, not for the deciduous dentition.

The relative age of monkeys' skulls has usually been estimated by the number and kind of teeth which have erupted (Krogman 1930a). The times at which either the

deciduous or permanent teeth erupt in the green monkey have not been recorded, and estimates of age based upon the picture of eruption must therefore be relative.

Schultz (1935) has shown by a study of 100 skulls that the sequence of tooth eruption in three genera of the subfamily Lasiopyginae ($=$Cercopithecinae) is practically identical. Krogman (1930b) has suggested various stages of dental eruption to delimit arbitrary age groups. Although a preliminary examination of both series of green monkeys' skulls used in the present study showed that the general order of tooth eruption was similar to that described for other Old World monkeys by Krogman (1930b) and Schultz (1935), intermediate stages between the completion of the deciduous and permanent dentitions could not be treated separately because the number of specimens representing these stages, whose sexes could be determined with reasonable certainty, was too few. All skulls in which all the permanent teeth had not erupted were consequently grouped together as subadults.

Schultz (1935) has observed that the wear of the permanent teeth and the closure of cranial sutures are not highly correlated in adult monkeys. In the present study, therefore, the relative ages of adult skulls possessing a full permanent dentition were assessed entirely on the basis of tooth wear. As some cusps of the first permanent molars may show signs of wear before the complete permanent dentition is in place (Schultz 1935), the degree of wear of the second lower permanent molars, which erupt relatively much later, was taken as an index of age.

Observation showed that the dentine of the protoconid and hypoconid of the lower molars, that is to say of the two lateral cusps, becomes exposed before that of the other cusps. All skulls in which the complete permanent dentition was present, and in which the dentine was not exposed on more than two cusps of the second lower molars, were consequently classified as young adults, and those with the dentine of the lower second molar exposed on more than two cusps, as old adults.

For the study of the permanent teeth, the following age groups were thus recognized:

(a) Subadults, including all stages between the eruption of the first and third permanent molars.

(b) Young adults with the complete permanent dentition in place, but with attrition on not more than two cusps of the lower second molars.

(c) Old adults with attrition on more than two cusps of the lower second molars.

Adequate samples of these age groups, in the St Kitts and African series, were available only for subadult females, young adult females, young adult males, and old adult males. Since comparable age groups for the two series were not available, measurements of the permanent teeth of subadult males and old adult females were not included in our final statistical analysis of the size of the teeth.

Statistical analysis

Except where otherwise stated, differences between means or variances which gave values of P less than or equal to 0·02 were regarded as significant.

(a) Computation of the basic statistics

The basic statistics for the series of measurements of teeth from the right and left side of each jaw were first computed separately and compared as described for anthropoid teeth by Ashton & Zuckerman (1950). In the total series of measurements of the St Kitts monkeys, there were only six significant differences between variances of corresponding measurements of right and left teeth, and none between their means. In the African group, there were only two differences between the variances, and none between the means. Even if both samples had been randomly drawn from one homogeneous population, with the level of significance arbitrarily fixed at 0·02, one in fifty would be expected to show a significant difference in any such series of multiple tests. In view of the very small numbers of significant differences which appeared in this analysis, and on the general grounds discussed elsewhere (Ashton & Zuckerman 1950), it was not considered justifiable to treat corresponding measurements of teeth from the left and right sides as independent variates.

The variance of the distribution of individual measurements was therefore analyzed so as to eliminate the variance 'within' monkeys, due to differences between the teeth on the two sizes of the jaw, while retaining any additional information which had been gained by measuring teeth from the two sides independently (Ashton & Zuckerman 1950).

(b) Comparison between different age groups

So as to discover whether the dimensions of the teeth did show any significant differences with age, the means and variances of the grouped data for both the subadult and young adult females, and for the young adult and old adult males of both African and St Kitts populations were compared by the same methods used in the comparisons of the teeth of the two sides of each jaw. It transpired that there were few differences between the tooth measurements of the subadult and young adult females, and between those of the cheek teeth of the young adult and old adult males. In these cases, therefore, means and variances were recomputed from all the data, irrespective of age.

(c) Comparisons between the African and St Kitts monkeys

Comparisons between the St Kitts and African monkeys were always made from statistics derived from this final resorting of the data.

Comparison between individual variates. The means and variances of corresponding attributes of the mainland and island stocks were compared by the methods already described (Ashton & Zuckerman 1950). These results are presented in tables 1 to 5.

The fiducial probabilities resulting from these *t* tests only give a measure of the confidence which can be placed in the observed differences between the two populations not being due to random sampling errors. They give no estimate whatever of the degree of divergence between the St Kitts and African stocks.

Sum of differences between the populations. The number of variates in each age group whose means and/or variances were greater or smaller in the African as

TABLES 1 TO 6. COMPARISONS BETWEEN THE DENTAL DIMENSIONS OF THE AFRICAN AND ST KITTS GREEN MONKEY

'+' indicates that the mean (or variance) of the African group is greater than that of the St Kitts.

'−' indicates that the mean (or variance) of the African group is less than that of the St Kitts.

All measurements are expressed in 10^{-2} cm.

"Significant" differences $P \leqslant 0.02$ are indicated in bold type. Italics are used where P falls in the range 0·05—0·02.

TABLE 1. COMPARISON BETWEEN THE DIMENSIONS OF THE MILK TEETH OF THE AFRICAN AND ST KITTS GREEN MONKEY

tooth	variable	African green monkey			St Kitts green monkey			comparison between African and St Kitts green monkey		
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio	difference between means	P
upper first incisor	max. labial transverse	47·53	1·69	9	48·00	0·87	4	+ 8·05	− 0·47	0·9 −0·8
	labial height	33·47	2·38	9	39·00	1·70	4	+ 4·16	− 5·53	0·2 −0·1
	max. lingual transverse	46·80	1·67	9	44·71	1·65	4	+ 2·21	+ 2·09	0·5 −0·4
	lingual height	40·47	1·58	9	41·43	2·18	4	+ 1·13	− 0·96	0·8 −0·7
	labio-lingual breadth	28·13	0·77	9	30·00	0·44	4	+ 6·68	− 1·87	0·2 −0·1
upper second incisor	max. labial transverse	34·89	0·92	10	38·25	0·76	4	+ 3·28	− 3·36	*0·05 −0·02*
	labial height	34·44	2·01	10	37·00	1·47	4	+ 4·18	− 2·56	0·5 −0·4
	max. lingual transverse	34·28	1·14	10	38·25	0·65	4	+ 6·97	− 3·97	*0·05 −0·02*
	lingual height	40·44	0·96	10	42·00	1·77	4	− 1·50	− 1·56	0·5 −0·4
	labio-lingual breadth	26·22	0·75	10	26·00	0·41	4	+ 7·68	+ 0·22	0·9 −0·8
lower first incisor	max. labial transverse	30·12	0·95	9	29·88	1·14	4	+ 1·49	+ 0·24	0·9 −0·8
	labial height	34·88	2·00	9	39·25	1·31	4	+ 4·96	− 4·37	0·2 −0·1
	max. lingual transverse	28·24	1·18	9	26·88	0·65	4	+ 7·13	+ 1·36	0·5 −0·4
	lingual height	39·82	2·43	9	43·88	0·68	4	+ 27·25	− 4·06	0·3 −0·2
	labio-lingual breadth	27·65	0·90	9	29·38	0·87	4	+ 2·31	− 1·73	0·3 −0·2
lower second incisor	max. labial transverse	29·78	1·12	9	32·50	1·08	5	+ 2·42	− 2·72	0·2 −0·1
	labial height	34·78	2·28	9	42·00	1·54	5	+ 4·93	− 7·22	0·1 −0·05
	max. lingual transverse	27·50	1·05	9	28·75	0·40	5	+ **15·80**	− 1·25	0·3 −0·2
	lingual height	41·11	2·34	9	46·38	2·29	5	+ 2·34	− 5·27	0·2 −0·1
	labio-lingual breadth	26·78	0·75	9	28·63	0·31	5	+ 13·33	− 1·85	0·2 −0·1

upper canine	labial height	49·85	1·45	14	54·52	1·15	14	1·68	+ −	4·67	**0·02-0·01**
	lingual height	50·31	1·33	14	53·36	1·10	14	1·51	+ −	3·05	0·1 −0·05
	max. A.-P. dimension	47·15	0·98	14	49·72	0·74	14	1·83	+ −	2·57	*0·05-0·02*
	labio-lingual breadth	28·92	0·74	14	30·40	0·51	14	2·19	+ −	1·48	0·2 −0·1
lower canine	labial height	46·39	1·27	12	51·96	0·84	14	2·13	+ −	5·57	≒**0·001**
	lingual height	46·09	1·11	12	48·44	0·92	14	1·33	+ −	2·35	0·2 −0·1
	max. A.-P. dimension	37·57	1·46	12	42·52	0·93	14	2·26	+ −	4·95	**0·01-0·001**
	labio-lingual breadth	26·13	1·10	12	25·92	0·62	14	2·86	+ +	0·21	0·9 −0·8
upper first molar	A.-P. length	51·23	0·87	13	53·58	0·54	14	2·62	+ −	2·35	*0·05-0·02*
	maximum breadth	39·27	0·96	13	39·50	0·46	14	4·35	+ −	0·23	0·9 −0·8
	index	76·65	1·43	13	73·81	0·75	14	3·61	+ +	2·84	0·1 −0·05
upper second molar	A.-P. length	54·27	1·09	13	56·68	0·62	13	3·17	+ −	2·41	0·1 −0·05
	trigone breadth	45·35	0·97	13	46·68	0·38	13	6·89	+ −	1·33	0·3 −0·2
	talon breadth	43·23	0·75	13	43·72	0·41	13	3·50	+ −	0·49	0·6 −0·5
	maximum breadth	45·50	0·94	13	46·68	0·38	13	6·45	+ −	1·18	0·3 −0·2
	trigone index	83·65	1·37	13	82·40	0·81	13	2·95	+ +	1·25	0·5 −0·4
	talon index	79·88	1·12	13	77·32	0·94	13	1·48	+ +	2·56	0·1 −0·05
lower first molar	A.-P. length	57·54	0·82	12	57·03	0·72	16	1·05	+ −	0·39	0·8 −0·7
	maximum breadth	31·13	0·59	12	31·20	0·33	16	2·53	+ +	0·07	1·0 −0·9
	index	54·17	0·78	12	53·93	0·86	16	1·51	− +	0·24	0·9 −0·8
lower second molar	A.-P. length	57·38	0·80	11	58·60	0·70	15	1·07	− −	1·22	0·3 −0·2
	trigonid breadth	37·24	0·81	11	39·57	0·44	15	2·37	+ −	2·33	**0·02-0·01**
	talonid breadth	39·14	0·89	11	39·00	0·47	15	2·49	+ +	0·14	0·9 −0·8
	maximum breadth	39·29	0·84	11	40·00	0·42	15	2·74	+ −	0·71	0·5 −0·4
	trigonid index	64·86	0·81	11	67·57	0·67	15	1·03	+ −	2·71	**0·02-0·01**
	talonid index	68·10	0·71	11	66·63	1·03	15	3·02	− +	1·47	0·3 −0·2

TABLE 2. COMPARISONS BETWEEN THE DIMENSIONS OF THE PERMANENT TEETH OF FEMALE AFRICAN AND ST KITTS GREEN MONKEY

tooth	variable	African green monkey — mean	standard error of mean	no. of animals	St Kitts green monkey — mean	standard error of mean	no. of animals	comparisons — variance ratio		Difference between means		P
upper first incisor	max. labial transverse	57·73	1·06	14	61·17	0·65	24	1·44	+	3·44	−	0·01–0·001
	labial height	84·81	2·68	14	83·08	2·74	24	1·93	−	1·73	+	0·7–0·6
	max. lingual transverse	55·69	1·46	14	60·00	0·94	24	1·29	+	4·31	−	0·02–0·01
	lingual height	69·19	1·77	14	70·56	1·81	24	1·93	−	1·37	−	0·7–0·6
	labio-lingual breadth	44·85	0·64	14	47·00	0·54	24	1·33	−	2·15	−	0·02–0·01
upper second incisor	max. labial transverse	40·20	0·96	13	41·84	0·70	22	1·10	+	1·64	−	0·2–0·1
	labial height	72·36	2·01	13	71·88	1·80	22	1·39	−	0·48	+	0·9–0·8
	max. lingual transverse	38·68	0·69	13	39·05	1·04	22	3·90	−	0·37	−	0·8–0·7
	lingual height	61·92	1·06	13	62·51	1·44	22	3·20	−	0·59	−	0·8–0·7
	labio-lingual breadth	39·20	0·83	13	39·84	0·57	22	1·23	+	0·64	−	0·6–0·5
lower first incisor	max. labial transverse	40·38	0·85	13	41·37	0·54	22	1·49	+	0·99	−	0·4–0·3
	labial height	83·12	2·11	13	82·09	2·18	22	1·76	−	1·03	+	0·8–0·7
	max. lingual transverse	39·38	1·03	13	39·77	0·72	22	1·23	+	0·39	−	0·8–0·7
	lingual height	68·50	1·88	13	72·74	1·35	22	1·17	+	4·24	−	0·1–0·05
	labio-lingual breadth	43·62	0·98	13	46·37	0·75	22	1·02	+	2·75	−	0·05–0·02
lower second incisor	max. labial transverse	37·76	0·68	13	38·86	0·76	22	2·11	−	1·10	−	0·4–0·3
	labial height	77·40	1·80	13	79·38	1·74	22	1·56	−	1·98	−	0·5–0·4
	max. lingual transverse	35·88	0·57	13	35·26	0·75	22	2·98	−	0·62	+	0·6–0·5
	lingual height	69·28	1·68	13	73·93	1·56	22	1·44	−	4·65	−	0·1–0·05
	labio-lingual breadth	42·72	1·20	13	45·31	0·56	22	2·71	+	2·59	−	0·05–0·02
upper canine	labial height	131·56	13·26	5	154·30	3·89	5	10·45	+	22·74	−	0·2–0·1
	lingual height	127·00	11·28	5	147·70	4·07	5	6·93	+	20·70	−	0·2–0·1
	max. A.-P. dimension	65·11	2·25	5	73·80	0·97	5	4·79	+	8·69	−	0·01–0·001
	labio-lingual breadth	43·67	0·33	5	46·20	1·48	5	21·75	−	2·53	−	0·2–0·1
lower canine	labial height	96·13	8·50	5	106·00	2·97	9	3·86	+	9·87	−	0·3–0·2
	lingual height	88·00	6·11	5	100·59	2·72	9	2·38	+	12·59	−	0·1–0·05
	max. A.-P. dimension	34·75	1·19	5	39·65	0·88	9	1·17	+	4·90	−	0·01–0·001
	labio-lingual breadth	53·25	1·38	5	59·24	1·07	9	1·29	−	5·99	−	0·01–0·001

upper first premolar	A.-P. length	43·29	1·01	12	45·26	0·78	19	1·17	+	1·97	0·2 −0·1
	maximum breadth	41·71	1·39	12	43·74	0·70	19	2·66	+	2·03	0·2 −0·1
upper second premolar	A.-P. length	44·87	0·99	12	48·10	0·65	20	1·40	+	3·23	0·01−0·001
	maximum breadth	48·61	0·70	12	51·33	0·44	20	1·49	+	2·72	0·01−0·001
lower first premolar	A.-P. length	55·32	1·41	11	57·13	1·20	16	1·02	−	1·81	0·4 −0·3
	maximum breadth	31·86	1·09	11	33·19	0·73	16	1·59	+	1·33	0·4 −0·3
lower second premolar	A.-P. length	52·67	1·23	12	55·13	0·79	16	1·87	+	2·46	0·1 −0·05
	maximum breadth	35·25	0·59	12	39·52	1·39	16	7·12	−	4·27	0·01−0·001
upper first molar	A.-P. length	59·56	0·90	17	62·62	0·67	26	1·18	+	3·06	0·01−0·001
	trigone breadth	52·41	0·79	17	56·10	0·42	26	2·36	+	3·69	<0·001
	talon breadth	50·15	0·70	17	52·63	0·47	26	1·43	+	2·48	0·01−0·001
	maximum breadth	52·41	0·79	17	56·10	0·42	26	2·36	+	3·69	<0·001
	trigone index	88·03	0·91	17	89·79	0·76	26	1·06	−	1·76	0·2 −0·1
	talon index	84·26	0·89	17	84·31	0·83	26	1·35	−	0·05	1·0 −0·9
upper second molar	A.-P. length	63·92	0·80	13	65·65	0·72	20	1·23	−	1·73	0·2 −0·1
	trigone breadth	60·73	1·04	13	64·58	0·58	20	2·10	+	3·85	0·01−0·001
	talon breadth	56·96	1·08	13	59·90	0·69	20	1·60	+	2·94	*0·05−0·02*
	maximum breadth	60·85	1·03	13	64·58	0·58	20	2·07	+	3·73	0·01−0·001
	trigone index	95·08	1·40	13	98·58	1·17	20	1·08	−	3·50	0·1 −0·05
	talon index	89·19	1·29	13	91·38	1·10	20	1·12	−	2·19	0·3 −0·2
upper third molar	A.-P. length	54·70	1·01	5	55·13	0·76	4	2·19	+	0·43	0·8 −0·7
	trigone breadth	53·50	1·01	5	59·25	0·61	4	3·42	+	5·75	0·01−0·001
	talon breadth	49·20	1·96	5	54·13	0·96	4	5·25	+	4·93	0·1 −0·05
	maximum breadth	53·50	1·01	5	59·25	0·61	4	3·42	+	5·75	0·01−0·001
	trigone index	97·80	1·60	5	107·63	1·34	4	1·78	+	9·83	0·01−0·001
	talon index	90·00	3·23	5	98·50	2·82	4	1·64	+	8·50	0·1 −0·05

E. H. Ashton and S. Zuckerman

TABLE 2. (contd.)

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey			
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio	difference between means		P
lower first molar	A.-P. length	61·21	0·98	15	63·06	0·67	25	1·29	+	1·85	0·2 –0·1
	trigonid breadth	43·90	0·56	15	46·22	0·46	25	1·12	–	2·32	**0·01–0·001**
	talonid breadth	45·10	0·74	15	46·43	0·54	25	1·10	+	1·33	0·2 –0·1
	maximum breadth	45·31	0·66	15	47·22	0·48	25	1·11	+	1·91	*0·05–0·02*
	trigonid index	71·79	0·59	15	73·51	0·74	25	2·63	–	1·72	0·2 –0·1
	talonid index	73·66	0·62	15	73·78	0·84	25	3·09	–	0·12	1·0 –0·9
lower second molar	A.-P. length	64·08	1·07	13	65·90	0·64	20	1·87	+	1·82	0·2 –0·1
	trigonid breadth	52·23	0·83	13	56·08	0·61	20	1·21	+	3·85	**<0·001**
	talonid breadth	52·31	0·90	13	53·21	0·57	20	1·65	+	0·90	0·4 –0·3
	maximum breadth	53·04	0·88	13	56·13	0·59	20	1·48	+	3·09	**0·01–0·001**
	trigonid index	81·65	0·96	13	85·26	0·98	20	1·55	–	3·61	**0·02–0·01**
	talonid index	81·81	1·21	13	80·85	1·10	20	1·23	–+	0·96	0·6 –0·5
lower third molar	A.-P. length	60·30	0·35	5	62·29	1·14	8	14·51	–	1·99	0·2 –0·1
	trigonid breadth	51·80	1·40	5	55·00	0·80	8	2·17	+	3·20	0·1 –0·05
	talonid breadth	44·10	1·33	5	45·71	1·29	8	1·32	–	1·61	0·5 –0·4
	maximum breadth	51·80	1·40	5	55·00	0·80	8	2·17	+	3·20	0·1 –0·05
	trigonid index	86·00	1·88	5	88·43	1·52	8	1·10	+	2·43	0·4 –0·3
	talonid index	73·10	2·16	5	73·43	2·16	8	1·41	–	0·33	1·0 –0·9

TABLE 3. COMPARISONS BETWEEN THE DIMENSIONS OF THE PERMANENT TEETH OF THE YOUNG
ADULT MALE AFRICAN AND ST KITTS GREEN MONKEY

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey			
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio	difference between means		P
upper first incisor	max. labial transverse	60·29	0·89	14	62·09	1·05	24	2·32	− −	1·80	0·3 −0·2
	labial height	86·14	2·18	14	79·47	2·74	24	2·65	− +	6·67	≒0·1
	max. lingual transverse	57·93	0·70	14	60·00	1·02	24	3·58	− −	2·07	0·2 −0·1
	lingual height	75·36	2·50	14	67·96	2·02	24	1·10	− +	7·40	*0·05−0·02*
	labio-lingual breadth	46·04	0·65	14	47·72	0·67	24	1·80	− −	1·68	0·2 −0·1
upper second incisor	max. labial transverse	41·39	0·93	14	41·30	1·02	23	1·98	− +	0·09	1·0 −0·9
	labial height	72·36	1·62	14	69·65	2·44	23	3·74	− +	2·71	0·4 −0·3
	max. lingual transverse	40·89	1·05	14	38·83	0·87	23	1·13	− +	2·06	0·2 −0·1
	lingual height	66·64	1·58	14	62·54	1·86	23	2·29	− +	4·10	0·2 −0·1
	labio-lingual breadth	40·57	0·72	14	41·07	0·56	23	1·00	+ −	0·50	0·6 −0·5
lower first incisor	max. labial transverse	41·85	0·51	13	41·67	0·66	23	2·93	− +	0·18	0·9 −0·8
	labial height	83·42	2·89	13	82·00	1·98	23	1·20	+ +	1·42	0·7 −0·6
	max. lingual transverse	39·85	0·79	13	40·33	0·69	23	1·37	− −	0·48	0·7 −0·6
	lingual height	76·04	1·95	13	73·28	1·67	23	1·30	− +	2·76	0·4 −0·3
	labio-lingual breadth	45·62	1·35	13	46·15	0·83	23	1·49	+ −	0·53	0·8 −0·7
lower second incisor	max. labial transverse	38·73	1·18	13	36·27	1·07	24	1·52	− +	2·46	0·2 −0·1
	labial height	78·73	2·22	13	76·13	2·14	24	1·72	− +	2·60	0·5 −0·4
	max. lingual transverse	37·42	1·03	13	35·06	0·67	24	1·30	+ +	2·36	0·1 −0·05
	lingual height	73·46	1·50	13	72·44	1·75	24	2·51	− +	1·02	≒0·7
	labio-lingual breadth	43·54	1·01	13	45·15	0·68	24	1·19	+ −	1·61	0·2 −0·1
upper canine	labial height	185·65	9·15	9	198·16	5·17	18	1·66	+ −	12·51	0·3 −0·2
	lingual height	179·59	8·80	9	190·03	5·56	18	1·33	+ −	10·44	0·4 −0·3
	max. A.-P. dimension	85·41	4·45	9	87·28	1·79	18	3·29	+ −	1·87	0·7 −0·6
	labio-lingual breadth	53·41	2·83	9	59·16	1·71	18	1·45	+ −	5·75	0·1 −0·05
lower canine	labial height	139·15	5·37	13	145·75	3·25	18	1·98	+ −	6·60	0·3 −0·2
	lingual height	128·77	5·06	13	133·53	3·33	18	1·67	+ +	4·76	0·5 −0·4
	max. A.-P. dimension	47·08	1·58	13	45·69	1·05	18	1·62	+ +	1·39	0·5 −0·4
	labio-lingual breadth	72·31	2·83	13	73·92	1·65	18	2·14	+ −	1·61	0·7 −0·6

TABLE 4. COMPARISONS BETWEEN THE DIMENSIONS OF THE PERMANENT TEETH OF THE OLD ADULT MALE AFRICAN AND ST KITTS GREEN MONKEY

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey			
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio		difference between means	P
upper first incisor	max. labial transverse	56·33	3·17	3	59·07	1·31	16	1·18	+ −	2·74	0·5 −0·4
	labial height	78·50	2·47	3	50·17	3·45	16	9·78	− +	28·33	**0·01–0·001**
	max. lingual transverse	53·50	6·25	3	54·37	1·38	16	4·13	+ −	0·87	0·9 −0·8
	lingual height	68·50	5·20	3	55·67	1·45	16	2·57	+ +	12·83	**0·01–0·001**
	labio-lingual breadth	46·83	1·50	3	49·17	0·68	16	1·04	− −	2·34	0·2 −0·1
upper second incisor	max. labial transverse	34·57	2·94	4	39·71	0·67	15	4·77	+ −	5·14	**0·02–0·01**
	labial height	72·14	6·62	4	48·29	3·25	15	1·03	+ +	23·85	**0·01–0·001**
	max. lingual transverse	35·86	2·48	4	38·36	0·90	15	1·89	+ −	2·50	0·3 −0·2
	lingual height	67·14	5·51	4	54·18	1·72	15	2·57	+ +	12·96	**0·01–0·001**
	labio-lingual breadth	39·71	0·44	4	41·46	0·67	15	9·35	− −	1·75	0·3 −0·2
lower first incisor	max. labial transverse	33·60	1·65	5	37·91	1·23	18	1·90	− −	4·31	0·1 −0·05
	labial height	48·80	4·30	5	54·97	2·95	18	1·59	− −	6·17	0·4 −0·3
	max. lingual transverse	32·50	3·14	5	34·21	1·54	18	1·22	+ −	1·71	0·7 −0·6
	lingual height	54·10	4·66	5	65·71	2·32	18	1·19	+ −	11·61	*0·05–0·02*
	labio-lingual breadth	46·90	1·14	5	46·85	0·63	18	1·04	− +	0·05	1·0 −0·9
lower second incisor	max. labial transverse	30·44	2·21	5	32·84	1·30	18	1·19	− −	2·40	0·4 −0·3
	labial height	51·89	2·93	5	53·48	2·93	18	3·44	− −	1·59	0·8 −0·7
	max. lingual transverse	31·11	1·92	5	30·81	1·47	18	2·02	− +	0·30	1·0 −0·9
	lingual height	52·22	3·60	5	65·16	2·39	18	1·52	− −	12·94	**0·02–0·01**
	labio-lingual breadth	44·11	0·93	5	47·00	0·70	18	1·94	− −	2·89	0·1 −0·05
upper canine	labial height	171·00	13·59	5	173·56	6·06	19	1·33	+ −	2·56	0·9 −0·8
	lingual height	162·67	9·85	5	172·41	6·47	19	1·63	− −	9·74	0·5 −0·4
	max. A.-P. dimension	79·00	4·55	5	87·59	1·41	19	2·75	+ −	8·59	*0·05–0·02*
	labio-lingual breadth	56·00	2·35	5	61·68	1·27	19	1·10	− −	5·68	*≒0·05*
lower canine	labial height	131·20	7·35	6	118·29	6·48	17	2·41	− +	12·91	*≒0·3*
	lingual height	120·70	7·31	6	114·23	5·47	17	1·74	− +	6·47	0·6 −0·5
	max. A.-P. dimension	44·80	1·36	6	45·61	0·83	17	1·17	− −	0·81	0·7 −0·6
	labio-lingual breadth	69·50	2·67	6	72·97	1·36	17	1·24	+ −	3·47	0·3 −0·2

TABLE 5. COMPARISONS BETWEEN THE DIMENSIONS OF THE PERMANENT TEETH OF THE ADULT MALE AFRICAN AND ST KITTS GREEN MONKEY

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey			
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio		difference between means	P
upper first premolar	A.-P. length	44·71	0·84	21	46·49	0·48	50	1·29	+ −	1·78	0·1 –0·05
	maximum breadth	44·39	0·78	21	44·52	0·44	50	1·34	+ −	0·13	0·9 –0·8
upper second premolar	A.-P. length	46·76	0·68	21	48·71	0·41	51	1·19	+ −	1·95	**0·02–0·01**
	maximum breadth	49·40	0·60	21	52·15	0·38	51	1·05	+ −	2·75	**<0·001**
lower first premolar	A.-P. length	65·46	1·47	19	62·93	0·81	49	1·29	+ +	2·53	0·2 –0·1
	maximum breadth	33·97	0·49	19	35·00	0·40	49	1·73	− −	1·03	0·2 –0·1
lower second premolar	A.-P. length	55·21	0·61	20	56·64	0·56	49	2·21	− −	1·43	0·2 –0·1
	maximum breadth	35·89	0·49	20	37·49	0·46	49	2·21	− −	1·60	*≑0·05*
upper first molar	A.-P. length	62·05	0·63	22	63·23	0·42	50	1·01	+ −	1·18	0·2 –0·1
	trigone breadth	56·53	0·75	22	58·09	0·42	50	1·37	+ −	1·56	0·1 –0·05
	talon breadth	53·88	0·74	22	55·05	0·39	50	1·60	+ −	1·17	0·2 –0·1
	maximum breadth	56·81	0·75	22	58·19	0·43	50	1·33	+ −	1·38	0·1 –0·05
	trigone index	91·26	1·18	22	91·96	0·68	50	1·33	+ −	0·70	0·6 –0·5
	talon index	86·98	1·00	22	87·12	0·54	50	1·50	+ −	0·14	≑0·9
upper second molar	A.-P. length	64·98	0·69	22	66·24	0·40	51	1·29	+ −	1·26	0·2 –0·1
	trigone breadth	63·91	0·85	22	66·70	0·40	51	1·94	+ −	2·79	**0·01–0·001**
	talon breadth	59·63	0·95	22	62·15	0·44	51	2·02	+ −	2·52	**0·01–0·001**
	maximum breadth	64·21	0·86	22	66·79	0·38	51	2·19	+ −	2·58	**0·01–0·001**
	trigone index	98·42	1·07	22	100·85	0·64	51	1·21	+ −	2·43	*0·05 –0·02*
	talon index	91·98	1·60	22	93·90	0·58	51	3·28	+ −	1·92	0·3 –0·2

TABLE 5. (contd.)

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey		
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio	difference between means	P
upper third molar	A.-P. length	61·11	0·94	19	60·91	0·53	43	1·42 + +	0·20	0·9 –0·8
	trigone breadth	59·84	1·07	19	62·27	0·40	43	3·19 + –	2·43	0·05–0·02
	talon breadth	53·35	1·06	19	56·70	0·55	43	1·68 + –	3·35	0·01–0·001
	maximum breadth	59·84	1·07	19	62·30	0·40	43	3·26 + –	2·46	0·05–0·02
	trigone index	98·05	1·34	19	102·45	0·75	43	1·44 + –	4·40	0·01–0·001
	talon index	87·30	0·85	19	93·22	0·97	43	2·90 – –	5·92	<0·001
lower first molar	A.-P. length	62·28	0·76	20	63·26	0·50	47	1·03 – –	0·98	0·3 –0·2
	trigonid breadth	45·03	0·63	20	46·98	0·37	47	1·22 + –	1·95	0·01–0·001
	talonid breadth	46·87	0·73	20	47·48	0·34	47	2·01 + –	0·61	0·4 –0·3
	maximum breadth	47·03	0·65	20	47·91	0·32	47	1·81 + –	0·88	0·2 –0·1
	trigonid index	72·38	0·95	20	74·49	0·66	47	1·13 – –	2·11	0·1 –0·05
	talonid index	75·36	0·96	20	75·25	0·61	47	1·07 + +	0·11	1·0 –0·9
lower second molar	A.-P. length	65·80	0·80	20	66·73	0·44	50	1·34 + –	0·93	0·3 –0·2
	trigonid breadth	54·23	0·84	20	57·93	0·46	50	1·38 + –	3·70	<0·001
	talonid breadth	54·30	0·74	20	54·95	0·51	50	1·18 – –	0·65	0·5 –0·4
	maximum breadth	55·13	0·78	20	58·10	0·43	50	1·33 + –	2·97	<0·001
	trigonid index	82·43	0·93	20	86·86	0·61	50	1·05 – –	4·43	<0·001
	talonid index	82·63	0·95	20	82·38	0·73	50	1·45 – +	0·25	0·9 –0·8
lower third molar	A.-P. length	65·61	0·88	18	66·73	0·53	46	1·12 + –	1·12	0·3 –0·2
	trigonid breadth	55·14	1·10	18	57·50	0·56	46	1·61 + –	2·36	0·05–0·02
	talonid breadth	49·17	0·91	18	49·84	0·45	46	1·70 + –	0·67	0·5 –0·4
	maximum breadth	55·17	1·09	18	57·63	0·50	46	1·94 + –	2·46	0·05–0·02
	trigonid index	84·00	1·36	18	86·31	0·71	46	1·47 + –	2·31	0·2 –0·1
	talonid index	75·00	1·02	18	74·85	0·67	46	1·04 – +	0·15	1·0 –0·9

compared with the St Kitts population were counted separately, and the totals tested both for homogeneity and for deviation from equality by a χ^2 test described by Mather (1938, 1946). These results are summarized in tables 7 to 12, where the data are assembled according to the pooled or separate comparisons made previously, and where the incisors and canines are dealt with as one group and the cheek teeth as another.

RESULTS

The morphological pattern of the teeth is the same in the island and mainland stocks of green monkey. The quantitative analysis undertaken in the present study shows, however, that the two populations differ significantly in thirteen out of 188 variances, and thirty-nine out of 188 means of the dimensions compared (tables 1 to 5). The number of significant differences in both variances and means is significantly greater than the 1 in 50 which a corresponding series of multiple tests applied to pairs of random samples from a single homogeneous population would be expected to show, if the level of significance were fixed at $P = 0.02$ (table 6). Real differences therefore exist between the two populations. Differences additional to those demonstrated may, in fact, exist between some of the other attributes which were examined, but they did not emerge in the analysis of the samples which were available.

Comparison between variances

(a) Linear measurements of incisors and canines

Of the total of 112 measurements of the incisors and canines, the African population is more variable than the St Kitts in sixty-six, and less so in forty-six (table 7). The insignificant value of the deviation χ^2 implies that these figures do not indicate any significant trend towards a change in overall variability between the two populations. The four age groups are markedly heterogeneous, and when considered separately, only the measurements of the milk dentition show any significant trend, which is towards a decrease in the variability of the St Kitts monkeys.

(b) Linear measurements of premolars and molars

Only twelve of the total of seventy-six linear measurements of the cheek teeth of the St Kitts monkeys are more variable than those of the African population (table 8). This indicates a highly significant trend towards a decrease in overall variability of the St Kitts population compared with the African. The groups are homogeneous, and the trend is obvious in each.

(c) Molar indices

Analysis does not reveal any significant trend towards a change in variability of the indices of the molar teeth either when taken together or when taken separately in age groups, which are homogeneous. Of the total of thirty examined, the African population is more variable in fourteen and the St Kitts in sixteen (table 9).

TABLE 6. COMPARISON BETWEEN THE TOTAL NUMBERS OF SIGNIFICANT $(P \leqslant 0.02)$ AND NON-SIGNIFICANT DIFFERENCES BETWEEN THE MEANS AND VARIANCES OF LINEAR MEASUREMENTS OF THE TEETH OF THE AFRICAN AND ST KITTS POPULATIONS

Each value of χ^2 tests for deviation from a 1 : 49 ratio.

	no. of significant differences	no. of non-significant differences	χ^2	P
variances	13	175	23·17 [1]	<0·001
means	39	149	337·02 [1]	<0·001

Comparison between means

(a) Linear measurements of incisors and canines

The St Kitts population is larger than the African in eighty-one out of the total of 112 incisor and canine dimensions studied (table 10). This deviation indicates a highly significant trend towards an increase in size of the St Kitts population. The groups are, however, heterogeneous, although not markedly so. If they are treated as such, and the appropriate test applied, the indication of an increase in size becomes less reliable. When, however, each age group is considered separately, significant trends towards an increase in the size of the St Kitts population are seen in the measurements of the deciduous teeth and in those of the permanent teeth of the grouped females. Significant trends are not revealed in the teeth of the young or old adult males.

(b) Linear measurements of premolars and molars

There is a highly significant trend towards an increase in size of the St Kitts monkeys as compared with the African in each of the three groups in which these measurements are compared (table 11). In only three of the total of seventy-six measurements are the St Kitts monkeys smaller. The total deviation is highly significant, and the three groups are homogeneous.

(c) Molar indices

Although the molar indices of the St Kitts monkeys are greater than those of the African in twenty-one out of a total of thirty compared, the three groups are markedly heterogeneous, and when account is taken of this, the total deviation is not significant (table 12). When the groups are considered separately, a significant trend towards an increase in the molar indices of the St Kitts monkeys is apparent in the grouped females, but not in either the grouped males or in the deciduous dentition.

Hence, although there has been no change in their shape, the premolar and molar teeth of the St Kitts monkeys are larger and less variable than those of the present-day African stock. The results are not so clearly defined in the case of the incisors and canines, although there are indications that these teeth may also be larger in the St Kitts population.

The quantities in the columns headed + are the numbers of attributes in which the African green monkey is larger than the St Kitts.

Each χ^2 is a measure of the deviation of the group from equality.

The figures in square brackets are the number of degrees of freedom for each value of χ^2.

TABLE 7. OVERALL COMPARISON BETWEEN THE VARIANCES OF MEASUREMENTS
OF THE INCISOR AND CANINE TEETH OF THE AFRICAN AND
ST KITTS GREEN MONKEY

group	+	−	χ^2	P
milk dentition	27	1	24·14 [1]	<0·001
grouped females	14	14	0·00 [1]	1·00
young adult males	13	15	0·14 [1]	0·8–0·7
old adult males	12	16	0·57 [1]	0·5–0·3
total χ^2			24·85 [4]	
total numbers of observations and deviation χ^2	66	46	3·571 [1]	0·1–0·05
heterogeneity χ^2			21·279 [3]	<0·001

mean square of deviation $\chi^2 = 3\cdot571$
mean square of heterogeneity $\chi^2 = 7\cdot09$
variance ratio <1

TABLE 8. OVERALL COMPARISON BETWEEN THE VARIANCES OF MEASUREMENTS
OF THE PREMOLAR AND MOLAR TEETH OF THE AFRICAN AND
ST KITTS GREEN MONKEY

group	+	−	χ^2	P
milk dentition	11	1	8·33 [1]	0·01–0·001
grouped females	26	6	12·50 [1]	<0·001
grouped males	27	5	15·13 [1]	<0·001
total χ^2			35·96 [3]	
total numbers of observations and deviation χ^2	64	12	35·58 [1]	<0·001
heterogeneity χ^2			0·38 [2]	0·9–0·8

TABLE 9. OVERALL COMPARISON BETWEEN THE VARIANCE OF THE MOLAR
INDICES OF THE ST KITTS AND AFRICAN GREEN MONKEY

group	+	−	χ^2	P
milk dentition	4	2	0·67 [1]	0·5–0·3
grouped females	3	9	3·00 [1]	0·1–0·05
grouped males	7	5	0·33 [1]	0·7–0·5
total χ^2			4·00 [2]	
total number of observations and deviation χ^2	14	16	0·13 [1]	0·8–0·7
heterogeneity χ^2			3·87 [2]	0·2–0·1

TABLE 10. OVERALL COMPARISON BETWEEN THE MEANS OF LINEAR MEASURE-
MENTS OF THE INCISOR AND CANINE TEETH OF THE ST KITTS
AND AFRICAN GREEN MONKEY

group	+	−	χ^2	P
milk dentition	5	23	11·57 [1]	<0·001
grouped females	4	24	14·29 [1]	<0·001
young adult males	14	14	0·00 [1]	1·00
old adult males	8	20	5·14 [1]	0·05–0·02
total χ^2			31·00 [4]	
total number of observations and deviation χ^2	31	81	22·32 [1]	<0·001
heterogeneity χ^2			8·68 [3]	0·05–0·02

mean square of deviation χ^2 = 22·32 [1]
mean square of heterogeneity χ^2 = 2·89 [3]
variance ratio = 7·72
P = 0·1–0·05

TABLE 11. OVERALL COMPARISON BETWEEN THE MEANS OF LINEAR MEASURE-
MENTS OF THE PREMOLAR AND MOLAR TEETH OF THE ST KITTS
AND AFRICAN GREEN MONKEY

group	+	−	χ^2	P
milk dentition	1	11	8·33 [1]	0·01–0·001
grouped females	0	32	32·00 [1]	<0·001
grouped males	2	30	24·50 [1]	<0·001
total χ^2			64·83 [3]	
total number of observations and deviation χ^2	3	73	64·47 [1]	<0·001
heterogeneity χ^2			0·36 [2]	0·9–0·8

TABLE 12. OVERAL COMPARISON BETWEEN THE MEANS OF THE MOLAR INDICES
OF THE ST KITTS AND AFRICAN GREEN MONKEY

group	+	−	χ^2	P
milk dentition	5	1	2·67 [1]	0·2 –0·1
grouped females	1	11	8·33 [1]	0·01–0·001
grouped males	3	9	3·00 [1]	0·1 –0·05
total χ^2			14·00 [3]	
total number of observations and deviation χ^2	9	21	4·80 [1]	0·05–0·02
heterogeneity χ^2			9·2 [2]	≃0·01

mean square of deviation χ^2 = 4·800 [1]
mean square of heterogeneity χ^2 = 4·60 [2]
variance ratio = 1·04
P > 0·20

Comparisons of pooled data

Although the skulls were divided into male and female groups according to accepted methods, it might be argued that the differences discussed above have emerged as a chance result of this division. To test this possibility, comparisons

were made between the St Kitts and mainland stocks without any division of the data according to sex. The comparison was limited to the thirty-two dimensions of the cheek teeth which were not affected by wear.

It was found that there were significant differences between the African and St Kitts monkeys in all the nineteen dimensions in which such differences ($P \leqslant 0.05$) had been observed when the data were subdivided (tables 13 and 14). Of the thirteen dimensions which did not reveal a significant difference in either sex of the two populations, five passed the 0.05 level of significance and three the 0.02 level when the data were pooled (table 15). The trend towards a significant increase in the mean and a decrease in the variance of the St Kitts stock is also still apparent when the comparisons are made between the pooled data (table 16).

DISCUSSION

It is clear that in the course of the estimated seventy to a hundred generations during which the animal has been isolated on St Kitts, the green monkey has diverged significantly in many dental dimensions from the present-day descendants of its parent African stock. These differences are real, in so far as the two samples which were examined cannot be regarded as being derived from one homogeneous population. The teeth of the island stock have become larger, and the molars and premolars less variable.

Colyer (1945, 1948 a, 1948 b) examined ninety-two of the ninety-five St Kitts skulls considered in this paper for dental irregularities. He records that in three of the seventy-six adults an extra tooth is present, and that in one, a single tooth has failed to appear; whereas of 104 cercopitheque skulls which he examined, only two had extra teeth. Variations in the roots of the third molars are present in 20.8 % (fifteen) of the seventy-two St Kitts skulls in which this tooth is fully or almost fully formed, and in three out of the thirty-six third molars in the control group. Positional variations are also more frequent in the teeth of the St Kitts monkeys. Thus four mandibular and one maxillary tooth are misplaced in these skulls, whereas only one of the controls shows a positional variation of a mandibular third molar. Colyer also records the presence of positional variations in the maxillary incisors of over a quarter of the St Kitts specimens; in one mandibular canine; and in sixteen premolars. Unfortunately, he does not give comparative data for his control series. On the basis of all these observations he concludes, however, that the St Kitts monkeys are more variable than the African in the discontinuously varying characters which he investigated.

His observations are summarized in table 17. It is apparent that when his three classes of abnormality—numerical variations, variations in the roots of the third molars, and positional variations of the third molars—are considered separately, the evidence is not quite adequate to show that the discrepancy between the African and the St Kitts populations is due to causes other than variation in random sampling from a single homogeneous population. Nevertheless, in each case the bias is in favour of an increased variability of the St Kitts population. If the total frequencies of occurrence of normal and abnormal conditions in each

TABLE 13. COMPARISON BETWEEN AFRICAN AND St KITTS MONKEY, USING THE POOLED DATA FOR MALES AND FEMALES FOR VARIATES IN WHICH THERE IS A SIGNIFICANT DIFFERENCE $(P \leqslant 0.05)$ BETWEEN THE POPULATIONS IN BOTH SEXES

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey			
			standard error of mean	no. of animals		standard error of mean	no. of animals			difference between means	
		mean			mean			variance ratio			P
upper second premolar	A.-P. length	46·09	0·58	33	48·54	0·34	71	1·33	+ −	2·45	<0·001
	maximum breadth	49·12	0·46	33	51·92	0·30	71	1·08	+ −	2·80	<0·001
lower second premolar	maximum breadth	35·65	0·38	32	37·98	0·49	65	3·49	− −	2·33	<0·001
upper second molar	trigone breadth	62·71	0·70	35	66·09	0·35	71	2·00	+ −	3·38	<0·001
	talon breadth	58·62	0·74	35	61·51	0·39	71	1·82	+ −	2·89	<0·001
	maximum breadth	62·94	0·71	35	66·16	0·34	71	2·18	+ −	3·22	<0·001
upper third molar	trigone breadth	58·49	1·02	24	62·00	0·39	47	3·53	+ −	3·51	0·01–0·001
	maximum breadth	58·49	1·02	24	62·03	0·39	47	3·59	+ −	3·54	0·01–0·001
lower first molar	trigonid breadth	44·54	0·44	35	46·71	0·29	72	1·09	+ −	2·17	<0·001
lower second molar	trigonid breadth	53·44	0·62	33	57·40	0·38	70	1·27	+ −	3·96	<0·001
	maximum breadth	54·30	0·60	33	57·54	0·37	70	1·32	+ −	3·24	<0·001

males and females grouped

TABLE 14. COMPARISON BETWEEN AFRICAN AND ST KITTS MONKEY, USING THE POOLED DATA FOR MALES AND FEMALES FOR VARIATES IN WHICH THERE IS A SIGNIFICANT DIFFERENCE ($P \leqslant 0.05$) BETWEEN THE POPULATIONS IN ONE SEX ONLY

tooth	variable	African green monkey			St Kitts green monkey			comparisons between African and St Kitts green monkey		
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals	variance ratio	difference between means	P
upper first molar	A.-P. length	60·95	0·56	39	63·02	0·36	76	1·27	+ 2·07	0·01–0·001
	trigone breadth	54·71	0·63	39	57·40	0·33	76	1·89	+ 2·69	<0·001
	talon breadth	52·23	0·59	39	54·21	0·33	76	1·67	+ 1·98	0·01–0·001
	maximum breadth	54·87	0·65	39	57·47	0·34	76	1·89	+ 2·60	<0·001
upper third molar	talon breadth	52·47	0·98	24	56·47	0·52	47	1·87	+ 4·00	<0·001
lower first molar	maximum breadth	46·29	0·48	35	47·67	0·27	72	1·59	+ 1·38	0·01–0·001
lower third molar	trigonid breadth	54·41	0·95	23	57·16	0·50	54	1·60	+ 2·75	0·01–0·001
	maximum breadth	54·43	0·94	23	57·26	0·46	54	1·89	+ 2·83	0·01–0·001

males and females grouped

TABLE 15. COMPARISON BETWEEN AFRICAN AND ST KITTS MONKEY, USING THE POOLED DATA FOR MALES AND FEMALES FOR VARIATES IN WHICH THERE IS NO SIGNIFICANT DIFFERENCE $(P \leqslant 0.05)$ BETWEEN THE POPULATIONS IN EITHER SEX

tooth	variable	Males and females grouped						comparisons between African and St Kitts green monkey		
		African green monkey			St Kitts green monkey			variance ratio	difference between means	P
		mean	standard error of mean	no. of animals	mean	standard error of mean	no. of animals			
upper first premolar	A.-P. length	44·18	0·65	33	46·17	0·41	69	1·23 +	1·99 −	0·01–0·001
	maximum breadth	43·40	0·73	33	44·31	0·37	69	1·91 +	0·91 −	0·3–0·2
lower first premolar	A.-P. length	61·68	1·39	30	61·50	0·74	65	1·64 +	0·18 +	1·0–0·9
	maximum breadth	33·19	0·53	30	34·56	0·36	65	1·01 +	1·37 −	0·05–0·02
lower second premolar	A.-P. length	54·23	0·63	32	56·27	0·47	65	1·17 −	2·04 −	0·02–0·01
upper second molar	A.-P. length	64·58	0·53	35	66·07	0·35	71	1·12 +	1·49 −	0·02–0·01
upper third molar	A.-P. length	59·74	0·94	24	60·40	0·55	47	1·56 +	0·66 −	0·6–0·5
lower first molar	A.-P. length	61·82	0·60	35	63·19	0·40	72	1·11 +	1·37 −	0·1–0·05
	talonid breadth	46·12	0·54	35	47·11	0·30	72	1·63 +	0·99 −	0·1–0·05
lower second molar	A.-P. length	65·12	0·65	33	66·49	0·36	70	1·53 +	1·37 −	0·05–0·02
	talonid breadth	53·52	0·59	33	54·45	0·41	70	1·01 −	0·93 −	0·2–0·1
lower third molar	A.-P. length	64·46	0·83	23	66·12	0·52	54	1·14 +	1·66 −	0·1–0·05
	talonid breadth	48·07	0·88	23	49·27	0·46	54	1·62 +	1·20 −	0·2–0·1

population are computed and compared, the St Kitts population, as Colyer suggests, is significantly more variable than the control.

TABLE 16. OVERALL COMPARISON BETWEEN BOTH MEANS AND VARIANCES OF LINEAR MEASUREMENTS OF THE PREMOLAR AND MOLAR TEETH OF THE ST KITTS AND AFRICAN GREEN MONKEY, USING THE POOLED DATA FROM MALES AND FEMALES

	+	−	χ^2	P
means	1	31	28·13 [1]	<0·001
variances	29	3	21·13 [1]	<0·001

TABLE 17. THE VARIABILITY OF SOME NON-QUANTITATIVE DENTAL CHARACTERS IN THE AFRICAN AND ST KITTS GREEN MONKEY
(ABSTRACTED FROM COLYER 1948*b*)

type of abnormality	African green monkey no. of abnormal observations	no. of normal observations	St Kitts green monkey no. of abnormal observations	no. of normal observations	χ^2	P
numerical variations	2	102	4	72	0·66	0·5–0·3
variations in the roots of the third molars	3	33	15	57	1·875	0·2–0·1
positional variations	1	103	5	67	2·986	0·1–0·05
totals	6	238	24	196	13·66	<0·001

In the dimensions of its teeth, therefore, the St Kitts population is less variable than the African, but in some 'non-quantitative' dental characters more so. In addition, the teeth of the St Kitts monkeys are larger than those of their parent stock. It is conceivable that the original breeding stock of monkey on St Kitts represented an extreme variant of its species, and that the characters now observed in its teeth are merely those of its progenitors. This, however, is improbable on the basis of pure chance, and is also inconsistent with the fact that the differences in variance between the island and mainland stocks are of a kind which on general selection theory might be expected to occur in an isolated group of animals which starts with the same genotypic constitution as its parent species. Assuming, as is almost certain, that change has occurred, there are two possible explanations—a direct effect of environment, or a genetical change in the island stock. The first is highly improbable; there is no indication that the teeth of any species of vertebrate are directly affected, in the way those of the St Kitts monkeys have been, by external factors acting during and after the process of eruption. A genetical explanation is the more likely, but is complicated by the apparent antithesis between the opposite trends of change in the variability of the 'quantitative' and 'non-quantitative' characters of the teeth. A way of explaining this antithesis is provided by the theory of polygenic inheritance (Mather 1943).

The teeth show continuous variation in the characters which we have measured. Thus any genetic change, however small, can express itself proportionately in the

measurements. Given constant heritable variation, its diminution, as a result of selection, will therefore be reflected as a corresponding reduction in phenotypic variation.

The characters investigated by Colyer, e.g. number of teeth present in the jaw, are, however, in a different category. They inevitably show discontinuities in their variation, and these discontinuities are unlikely to be related in any simple way to gene differences. The normal number of teeth may be produced by any of a wide range of genotypes differing polygenically among themselves, just as Sismanidis (1942) found to be the case with the scutellar bristles, normally four, in *Drosophila*. Only if the departure of the genotype from the mean balance exceeds some critical amount will the threshold be passed and an extra tooth be produced or a normal one fail to appear. Now if, in such a situation, selection should shift the mean balance of the genotype towards the critical threshold, a smaller departure from the new mean will produce the abnormality; and as the mean approaches the threshold, the frequency of abnormality will rise very sharply, other things being equal. Thus, even though the genetic variability is reduced by the selection which pushes the mean towards the threshold, the frequency of genotypes passing the threshold can rise, and with it, of course, the frequency of abnormality. In this way, reduction in genetic variability may easily be accompanied by an increase in phenotypic variation, where the character varies meristically as does number of teeth.

As already observed, it is safe to assume that evolution in the Old World Primates, as in the rest of the animal kingdom, is a measure of the effects of selection and isolation upon heritable variability. Isolation in the family Cercopithecidae has not depended upon the development of physiological barriers between its genera and species, as is shown by the fact that many interspecific and some intergeneric crosses have been recorded among them (Zuckerman 1933). On the other hand, the social habits of the Old World monkeys would seem to result in a breeding structure which, according to Mather's analysis of the general problem (1943), is favourable to a decrease in the free variability of polygenically controlled characters. In these animals intense competition usually occurs between the males for possession of the available females, each retaining the maximum number permitted by his position in the scale of dominance (Zuckerman 1932; Carpenter 1942). Although there is no direct evidence that the offspring of neighbouring groups do not interbreed, it is likely that mating between the offspring of different groups will be restricted rather than randomized, owing to the parental groups being kept apart by the sexual competition between the males.

Wright (1931, 1940, 1948) has deduced mathematically that division into small relatively inbreeding populations can favour the rapid spread of any genetic change even when opposed by adverse selection. In totally isolated groups the spread of new 'major' genes in such circumstances would frequently lead to extinction because such genes are usually deleterious. In partially isolated populations such as are encountered in the Old World Primates, however, any small local group which arrived at a genotype possessing a particular selective advantage would supply more than its share of new blood to other groups which would then

be upgraded by intergroup selection. Wright concludes that a partially isolated breeding structure would maintain the maximum amount of free variability, and would therefore be conducive to rapid evolutionary advance. Mather (1943), in discussing Wright's views, suggests that account should also be taken of two considerations based on the behaviour of polygenes. In the first place, fluctuations in the frequency of a single member of a polygenic system need not have a marked phenotypic effect, even if any increase in variability to which it contributed immediately became free. Secondly, such polygenic variability is potential rather than free, and its release depends on the arrangements of the polygenes in the chromosomes and on their recombination frequencies. Reduction in the size of the breeding unit would tend to decrease heterozygosity and so diminish the release of potential variability. This diminution would in turn tend to offset the free variability gained by gene drift in the small population. A small population would therefore tend to have less available variability, and hence less evolutionary plasticity. The results of the present study confirm such a loss of free variability in the St Kitts monkeys. At the same time there has been an evolutionary change in the mean values of the characters studied, and this may have occurred by selection operating on the group's original genetic constitution.

Further information about possible evolutionary changes in the St Kitts monkey is provided by Cameron (1930), who has studied the helminths which infect the intestinal tract of this animal. The worm of the genus *Subulura* which infests the St Kitts stock presents several small but constant differences, from the species *S. distans* which is found in the African green monkey, the most conspicuous of these being in the shape of the oesophagus. In Cameron's opinion the divergence between the two is sufficient to warrant the formation of the new species *S. neodistans* to accommodate the St Kitts worm. Cameron points out, however, that the St Kitts species may have been preserved in the isolated group after it had died out in the African, although he does not believe this likely.

What constitutes subspecific or specific divergence is largely an arbitrary matter, and one which, in the first place, only concerns taxonomic procedure. We propose to examine this question by estimating, by a series of discriminant functions (Fisher 1936), the degree of divergence between the teeth and skulls of the African and St Kitts green monkeys, and by comparing the result with corresponding estimates of the differences between recognized subspecies of the group *Cercopithecus aethiops* and species of *Cercopithecus*.

In the present paper we have been concerned primarily with the question of whether characters in one group of monkeys are significantly different in some definable attribute from the corresponding characters in another group. The problem is the same as that involved in comparing the skeletal characters of fossil apes and extant apes, except in so far as the comparison of two extant types does not presuppose that only one of the two has undergone evolutionary divergence. In comparisons in which fossils are involved it is idle to speculate about the possible ways the differences have arisen. For neither the morphologist nor the primate palaeontologist will be able to decide whether they are due to random changes within the genetic system, or to selection acting on free genetic variability. All

that can be said in such cases is that one attribute is significantly different, or not significantly different, from another, and that change has or has not occurred.

References

Ashton, E. H. & Zuckerman, S. 1950 *Phil. Trans.* B, **234**, 471.

Cameron, T. W. M. 1930 *J. Helminthol.* **8**, 49.

Campbell, — 1753 *Candid and impartial considerations on the nature of the sugar trade.* London: Baldwin.

Carpenter, C. R. 1942 *J. Comp. Psychol.* **23**, 113.

Colyer, J. F. 1945 *Dent. Rec.* **65**, 121.

Colyer, J. F. 1948a *Brit. Med. J.* p. 1202.

Colyer, J. F. 1948b *Proc. Roy. Soc. Med.* **41**, 845.

Doutt, J. K. 1942 *Ann. Carneg. Mus.* **29**, 61.

Elliott, D. G. 1912 *A review of the Primates*, 2. New York: American Mus. Nat. Hist. Monograph series, no. 1.

Fisher, R. A. 1936 *Ann. Eugen., Lond.,* **7**, 179.

Forbes, H. O. 1894 *A handbook of the Primates*, 2. London: Allen.

Jameson, H. L. 1898 *J. Linn. Soc. (Zool.),* **26**, 465.

Krogman, W. M. 1930a *Amer. J. Anat.* **46**, 303.

Krogman, W. M. 1930b *Amer. J. Anat.* **46**, 315.

Labat, J.-B. 1722 *Nouveau Voyage aux Isles de l'Amérique*, 1–6. Paris: Theodore le Gras.

Mather, K. 1938 *The measurement of linkage in heredity.* London: Methuen.

Mather, K. 1943 *Biol. Rev.* **18**, 32.

Mather, K. 1946 *Statistical analysis in biology*, 2nd ed. London: Methuen.

Montagu, M. F. Ashley 1937 *J. R. Anthrop. Inst.* **67**, 157.

Schlegel, H. 1876 *Mus. Pays-Bas, Simiae,* p. 73.

Schultz, A. H. 1935 *Amer. J. Phys. Anthrop.* **19**, 489.

Schwarz, E. 1928 *Ann. Mag. Nat. Hist.* (ser. 10), **1**, 649.

Sclater, P. L. 1866 *Proc. Zool. Soc. Lond.* p. 79.

Sclater, P. L. 1893 *Proc. Zool. Soc. Lond.* p. 243.

Simpson, G. G. 1944 *Tempo and mode in evolution.* New York: Columbia University Press.

Sismanidis, A. 1942 *J. Genet.* **44**, 204.

Wright, S. 1931 *Genetics,* **16**, 97.

Wright, S. 1940 *The new systematics*, p. 161. London: Oxford University Press.

Wright, S. 1948 *Evolution,* **2**, 279.

Zuckerman, S. 1926 *Proc. Zool. Soc. Lond.* p. 843.

Zuckerman, S. 1928 *Proc. Zool. Soc. Lond.* p. 1.

Zuckerman, S. 1932 *The social life of monkeys and apes.* London: Kegan Paul.

Zuckerman, S. 1933 *Functional affinities of man, monkeys and apes.* London: Kegan Paul.

The selective admission of antibodies to the foetus by the yolk-sac splanchnopleur in rabbits

By F. W. R. Brambell, F.R.S., W. A. Hemmings,
Megan Henderson and W. T. Rowlands

*Department of Zoology, University College of North Wales, and the
Veterinary Laboratory, Ministry of Agriculture, Bangor*

The yolk-sac splanchnopleur of 24-day rabbit embryos freely admits to the foetal circulation, antibodies prepared in the rabbit whereas antibodies prepared in cattle or horses are almost, but not entirely, excluded. This selectivity appears to be independent of the molecular dimensions of the antibody particles, for both rabbit agglutinins and haemolysins are admitted equally freely, whereas both equine antitoxins and bovine agglutinins are equally retarded. Selection appears to depend on the species-specific origin of the antibody molecules themselves, being independent of the other species-specific characters of the rabbit, bovine and equine sera employed.

The selectivity of the yolk-sac splanchnopleur as between rabbit and bovine agglutinins is relatively slight at 20 days post-coitum but develops progressively thereafter and is marked at 24 days post-coitum.

Individual embryos, even in the same litter, vary significantly in their capacity to admit ungulate antibodies.

The theoretical significance of the selectivity of the foetal splanchnopleur is discussed in relation to passive anaphylactic sensitization.

The technique of direct injection of immune sera into the lumen of the pregnant uterus was employed.

The antisera used throughout were the natural sera separated from the blood of immunized animals not having been subjected to any procedures for their refinement, purification or concentration.

INTRODUCTION

It has been shown (Brambell, Hemmings, Henderson, Parry & Rowlands 1949) that maternal circulating antibodies pass rapidly into the foetal circulation in rabbits during late pregnancy. They do so exclusively by way of the uterine lumen and the foetal yolk-sac splanchnopleur of the area vasculosa. Moreover, when immune serum of rabbits was introduced directly into the uterine lumen at 24 days of gestation the contained agglutinins were found to have attained high titres in the foetal serum within 24 hr. Only rabbit agglutinins were used in all these experiments on transmission through the yolk-sac splanchnopleur.

It was known already (Brambell, Hemmings & Rowlands 1948) that anti-brucella agglutinins of both rabbit and bovine origin could pass with apparently equal freedom from the maternal circulation through the bilaminar omphalopleur into the yolk-sac cavity of a blastocyst of 7 or 8 days of age. No differential selection of rabbit, as compared to bovine, antibody-globulin could be detected. It would be unjustifiable to assume that the yolk-sac splanchnopleur at 24 days would behave in a similar manner to the bilaminar omphalopleur at 8 days. Indeed, the pioneer work of Hartley (1948) suggested that the foetal membranes in late pregnancy are selective, for he showed that diphtheria antitoxin injected into the pregnant guinea-

pig passes to the offspring more readily if it is prepared in the guinea-pig than if it is prepared in the horse. Hence the experiments to be described were undertaken to determine if the yolk-sac splanchnopleur of the rabbit selects antibody globulins according to their specific origin.

The bilaminar omphalopleur and the yolk-sac splanchnopleur are morphologically quite distinct. The former consists of the continuous syncytial trophoblast on the outside and the extremely thin cellular entoderm on the inside. At scattered points on the surface of the blastocyst the trophoblast has penetrated through the uterine epithelium and invaded the vascular subepithelial tissues. Between these fusion areas the surface of the trophoblast is apposed to the uterine epithelium. The entoderm forms an extremely thin endothelium, which appears continuous but may be fenestrated, lining the inner surface of the trophoblast. The yolk-sac splanchnopleur, in contrast, consists of the entoderm on the outside, next the uterine lumen, and the splanchnic mesoderm of the area vasculosa on the inside. The inner surface of the splanchnic mesoderm is smooth and is covered by the thin squamous endothelium which lines the exocoel. The larger vitelline blood vessels project as prominent ridges on the outer surface of the mesoderm, which is rugose in consequence. The entoderm is cellular and forms a simple columnar epithelium covering the whole of the outer surface of the mesoderm including the projecting blood vessels. The surface of the entoderm is freely exposed to the uterine lumen and does not adhere to the uterine epithelium at any point. Hence different tissues are involved in the formation of the bilaminar omphalopleur and of the yolk-sac splanchnopleur, and the histological differentiation of the latter at 24 days is far in advance of that of the former at 8 days.

Experimental technique

The majority of the animals used were Dutch rabbits of the Agricultural Research Council's Compton strain of known history. A few animals of various breeds and unknown history were used to supplement these.

Laparotomies were performed under Nembutal followed by ether anaesthesia, using aseptic precautions. Adrenalin was administered during the operation in a few cases only, as an emergency measure when respiratory failure was feared. All intra-uterine injections were made into the tubal extremity of the uterine horn. Intravenous injections were made into ear veins.

Routine samples of maternal blood from an ear vein before operation, maternal heart blood immediately after killing, foetal amniotic fluid, heart blood and stomach contents were collected for titration. Samples of the immune sera injected were titrated at the same time, except in the first few experiments where reliance was placed on an initial titration of a sample from each bottle of immune serum.

All animals were killed by intravenous injection of magnesium chloride, a method which is humane and has the advantage of rendering the uterine musculature inactive.

The standard technique of agglutinin titration previously described (Brambell et al. 1948) was employed.

TABLE 1. INTRAVENOUS INJECTON OF BOVINE ANTI-BRUCELLA SERUM AT 22 DAYS POST-COITUM

no.	age at injection (days)	(hr.)	time of exposure (hr.)	dose of bovine serum (ml.)	(titre)	maternal titre after injection	maternal titre at killing	embryo 1	embryo 2	embryo 3	embryo 4	embryo 5	embryo 6	embryo 7	embryo 8
337	22	.	24	15	$\frac{1}{5120}$	.	$++\frac{1}{320}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
345	22	19	23	20	$\frac{1}{5120}$	$+\frac{1}{1280}$	$+++\frac{1}{320}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$
390	24	4	19	30	$\frac{1}{5120}$	.	$+\frac{1}{640}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
501	24	2	22	22	$\frac{1}{1280}$	$++\frac{1}{320}$	$++\frac{1}{40}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.	.	.

TABLE 2. INTRA-UTERINE INJECTION OF RABBIT OR BOVINE ANTI-BRUCELLA SERUM AT 24 DAYS POST-COITUM

no.	age at operation (days)	(hr.)	time of exposure (hr.)	dose of rabbit serum (ml.)	(titre)	dose of bovine serum (ml.)	(titre)	rabbit antibodies embryo 1	embryo 2	embryo 3	bovine antibodies embryo 1	embryo 2	embryo 3	embryo 4	embryo 5	embryo 6	embryo 7
152	24	0	24	7	$\frac{1}{2560}$	7	$\frac{1}{5120}$	$+\frac{1}{320}$	.	.	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.	.	.	.
132	23	19	23	7	$\frac{1}{2560}$	7	$\frac{1}{5120}$	$+++\frac{1}{160}$	$+++\frac{1}{160}$	$+++\frac{1}{160}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
348	24	0	24	7	$\frac{1}{2560}$	8	$\frac{1}{5120}$	$+\frac{1}{160}$	$+\frac{1}{80}$	.	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$++\frac{1}{20}$	$-\frac{1}{10}$
259	23	22	22	7	$\frac{1}{640}$	7	$\frac{1}{2560}$	$++\frac{1}{160}$	$++\frac{1}{160}$	$++\frac{1}{160}$	$-\frac{1}{10}$	$++\frac{1}{20}$	$-\frac{1}{10}$	$+\frac{1}{10}$	.	.	.
280	23	18	24	7	$\frac{1}{640}$	7	$\frac{1}{2560}$	$++\frac{1}{160}$	$+\frac{1}{160}$	.	$-\frac{1}{10}$	$+\frac{1}{20}$	$+++\frac{1}{20}$	$+\frac{1}{10}$	.	.	.
446	23	19	24	.	.	{7 / 6}	{$\frac{1}{5120}$ / $\frac{1}{5120}$}	.	.	.	$-\frac{1}{10}$ / $-\frac{1}{10}$	$-\frac{1}{10}$	$++\frac{1}{20}$	.	.	.	.
195	24	0	48	7	$\frac{1}{2560}$	7	$\frac{1}{5120}$	$+\frac{1}{320}$	$+\frac{1}{320}$	$+\frac{1}{320}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
443	23	20	49	.	.	4	$\frac{1}{2560}$	.	.	.	$-\frac{1}{10}$	.	.	.	.	.	.

TABLE 3. INTRAVENOUS INJECTION OF RABBIT ANTI-SHEEP SERUM AT 24 DAYS POST-COITUM

no.	age at injection (days)	time of exposure (hr.)	dose of rabbit serum (ml.)	(titre)	maternal titre after injection	embryo 1	embryo 2	embryo 3	embryo 4	embryo 5	embryo 6	embryo 7
289	24	22	29	.	$+++\frac{1}{5120}$	$+\frac{1}{640}$	$+\frac{1}{1280}$	$+\frac{1}{1280}$	$++\frac{1}{1280}$	$+++\frac{1}{640}$	$+++\frac{1}{1280}$	$++\frac{1}{1280}$

TABLE 4. INTRA-UTERINE INJECTION OF RABBIT ANTI-SHEEP AND BOVINE ANTI-BRUCELLA SERA AT 22 OR 24 DAYS POST-COITUM

no.	age at operation (days)	(hr.)	time of exposure (hr.)	dose of rabbit serum (ml.)	(titre)	dose of bovine serum (ml.)	(titre)	titres of rabbit antibodies in embryonic sera embryo 1	embryo 2	embryo 3	embryo 4	titres of bovine antibodies in embryonic sera embryo 1	embryo 2	embryo 3
401	24	0	3	7	$\frac{1}{1280}$	4	$\frac{1}{5120}$	$+\frac{1}{80}$	$+++\frac{1}{40}$	$+++\frac{1}{40}$	$+++\frac{1}{20}$	$-\frac{1}{10}$	.	.
308	21	22	46	7	$\frac{1}{2560}$	7	$\frac{1}{5120}$	$++\frac{1}{160}$	$++\frac{1}{320}$	$+\frac{1}{320}$	$+++\frac{1}{160}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$

TABLE 5. INTRA-UTERINE INJECTION OF MIXED BOVINE ANTI-BRUCELLA AND RABBIT ANTI-SHEEP SERUM AT 24 DAYS POST-COITUM

no.	age at operation (days)	(hr.)	time of exposure (hr.)	natural haemolysin titre of maternal serum	rabbit haemolysin or bovine agglutinin	dose and titre of mixed serum injected on right (ml.)	(titre)	embryo 1	embryo 2	embryo 3	embryo 4	embryo 5	embryo 6	dose and titre of mixed serum injected on left (ml.)	(titre)	embryo 1	embryo 2	embryo 3	embryo 4
1	24	0	24	$+\frac{1}{40}$	haemolysin	.	.	.	.	.	.	.	.	7	.	$+\frac{1}{640}$	$+++\frac{1}{640}$	.	.
					agglutinin	.	.	.	.	.	.	.	.	7	.	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
384	24	0	23	$-\frac{1}{10}$	haemolysin	6	$+++\frac{1}{1280}$	$+++\frac{1}{80}$	$+\frac{1}{80}$	.	.	.	.	6	$++\frac{1}{2560}$	$++\frac{1}{80}$	$+++\frac{1}{40}$	.	.
					agglutinin	6	$++\frac{1}{640}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.	.	.	6	$++\frac{1}{1280}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
420	23	22	23	$-\frac{1}{10}$	haemolysin	7	$+\frac{1}{640}$	$++\frac{1}{10}$	$++\frac{1}{40}$	$+\frac{1}{40}$	$+\frac{1}{40}$	$++\frac{1}{20}$	.	.	.	.	.
					agglutinin	7	$+++\frac{1}{1280}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$+++\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.	.	.	.
465	23	23	24	$-\frac{1}{10}$	haemolysin	8	$+\frac{1}{320}$	$+\frac{1}{40}$	$+\frac{1}{40}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	$+\frac{1}{20}$	$++\frac{1}{80}$	4	$+\frac{1}{320}$	$+\frac{1}{80}$	$++\frac{1}{40}$	.
					agglutinin	8	$+++\frac{1}{640}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	4	$+++\frac{1}{1280}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.
439	23	18	23	$-\frac{1}{10}$	haemolysin	6	$+\frac{1}{5120}$	$+++\frac{1}{160}$	$++\frac{1}{640}$	$+\frac{1}{320}$	.	.	.	4	$+\frac{1}{5120}$	$+++\frac{1}{640}$	$+++\frac{1}{160}$	.
					agglutinin	6	$+++\frac{1}{2560}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.	.	4	$++\frac{1}{2560}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.
497	23	19	23	$++\frac{1}{40}$	haemolysin	4	$+\frac{1}{2560}$	$-\frac{1}{10}$	$++\frac{1}{160}$	.	.	.	.	7·5	$++\frac{1}{10240}$	$+++\frac{1}{320}$	$-\frac{1}{10}$	$+\frac{1}{640}$	$+++\frac{1}{160}$
					agglutinin	4	$+\frac{1}{2560}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.	.	.	7·5	$++\frac{1}{1280}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$

No. 420 received an intravenous injection of 21 ml. bovine anti-brucella serum $+++\frac{1}{2560}$ 1 hr. after the intra-uterine injection, giving an immediate titre of $++\frac{1}{1280}$ and a final titre of $+++\frac{1}{80}$ in the maternal serum.
All haemolysin titrations were made with inactivated serum and perceptible reactions of less than one plus are not recorded in this table.

TABLE 6. INTRA-UTERINE INJECTION OF EQUINE ANTI-DIPHTHERIA AND RABBIT AND BOVINE ANTI-BRUCELLA SERA

no.	age at operation (days)	(hr.)	time of exposure (hr.)	antitoxin or agglutinin	maternal serum at killing	dose of equine antitoxin and/or rabbit agglutinin (ml.)	units/titre	embryo 1	embryo 2	embryo 3	embryo 4	embryo 5	dose of equine antitoxin and/or bovine agglutinin (ml.)	units/titre	embryo 1	embryo 2	embryo 3	embryo 4	embryo 5
258	23	18	24	agglutinin and antitoxin	0·1	6	$\frac{1}{2560}$	$+\frac{1}{320}$	$++\frac{1}{320}$	.	.	.	6	890	4	4	3	circa 2	5
215	23	22	24	agglutinin and antitoxin	0·2	7	$\frac{1}{2560}$	$+++\frac{1}{160}$	$+\frac{1}{80}$	$++\frac{1}{80}$	$+++\frac{1}{40}$	.	7	890	1	1	.	.	.
340	26	0	48	agglutinin and antitoxin	.	7	.	$+\frac{1}{320}$	$+++\frac{1}{80}$	$+++\frac{1}{80}$	.	.	7	890	5	2	5	.	.
267	26	0	48	agglutinin and antitoxin	1	4·5	.	$+\frac{1}{160}$	.	.	.	.	6	890	2	2	2	2	1
440*	23	3	45	antitoxin	0·17	.	400	0·083	0·14	.	.	.	.	400	0·083	0·083	.	.	.
			24	agglutinin	$-\frac{1}{10}$	.	$\frac{1}{640}$	$+\frac{1}{20}$	$+++\frac{1}{20}$	.	.	.	.	$\frac{1}{2560}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.	.
441	24	0	24	antitoxin	0·15	.	240	0·11	.	.	.	.	.	370	0·14	1·4	0·24	0·16	.
				agglutinin	$-\frac{1}{10}$	.	$\frac{1}{320}$	$++\frac{1}{20}$	.	.	.	.	.	$\frac{1}{1280}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$+\frac{1}{10}$	$-\frac{1}{10}$	.
311	23	20	23	antitoxin	0·3	.	310	0·14	0·125	0·28	0·067	0·1	.	320	0·125	.	.	.	.
				agglutinin	$-\frac{1}{10}$	.	$\frac{1}{320}$	$+\frac{1}{40}$	$+++\frac{1}{20}$	$+++\frac{1}{40}$	$+\frac{1}{40}$	$+\frac{2}{40}$	.	$\frac{1}{1280}$	$-\frac{1}{10}$	.	.	.	.
500	23	23	24	antitoxin	0·21	.	290	0·11	0·125	0·09	0·125	0·13	.	.	.	.	.	.	.
				agglutinin	$-\frac{1}{10}$	.	$\frac{1}{320}$	$+++\frac{1}{40}$	$+++\frac{1}{20}$	$+\frac{1}{40}$	$+\frac{1}{40}$	$+\frac{1}{40}$	.	.	.	.	.	.	.

* For explanation see text.

Complement used in the haemolysin titrations was either freshly collected guinea-pig serum or preserved serum (Burroughs Wellcome). In either case the tests were made up to contain in each dilution twice the amount of complement, as determined by previous titration, necessary for complete lysis of the antigen. The test samples were inactivated by heating to 56° C for 30 min. and serial dilutions of $\frac{1}{10}$ to $\frac{1}{5120}$ were made up. The antigen employed was a 2·5 % saline suspension of triple washed sheep red blood corpuscles packed by standard centrifugation. Tests were incubated for one hour at 37° C in a water-bath and read against a standard light. Evaluation was recorded as:

+ + + Complete haemolysis. Liquid quite transparent and clear.

  + + Incomplete haemolysis. Liquid almost transparent but with some opalescence due to residual cells.

   + Incomplete haemolysis. Liquid opalescent but still markedly more translucent than in negatives.

   — Red cells unaffected, unevenly distributed owing to sedimentation, leaving the supernatant clear and uncoloured.

Determinations of antitoxin concentration by intracutaneous injection of diphtheria toxin-sample mixtures into guinea-pigs were kindly undertaken on our behalf by Dr C. L. Oakley, of the Wellcome Physiological Research Laboratories, to whom we are greatly indebted for this assistance.

Rabbit anti-sheep haemolysin sera were prepared by a course of injections of 50 % saline suspension of red cells. Doses of 0·5 ml. were given intravenously on each of the first 3 days of the first week, 2·0 ml. intraperitoneally on the first day of the second week with 1·0 ml. intravenously on the two succeeding days, 3·0 ml. intraperitoneally on the first day of the third week with 1·5 ml. intravenously on the two succeeding days, this last dosage being repeated twice during two further weeks. Collection of serum followed 10 days after the last injection. In some cases after a rest period of 4 to 6 weeks doses of 3·0 and 10·0 ml. were given intraperitoneally with an interval of 7 days between them and serum was collected again 10 days after the last injection.

The bovine and rabbit anti-brucella sera used were supplied by the Ministry's Veterinary Research Laboratories at Weybridge, and the equine sera by the Lister Institute of Preventive Medicine and by the Wellcome Physiological Research Laboratories.

We are greatly indebted to all these institutions for these supplies.

RESULTS

Intravenous injection of bovine anti-brucella serum

Bovine anti-brucella serum was administered intravenously to four rabbits on the 23rd to 25th days of pregnancy. The animals were killed after *c*. 24 hr. No agglutinins could be detected in the sera of any of the 24 embryos at a titre of $\frac{1}{10}$, as can be seen from the data summarized in table 1.* Titres of at least $\frac{1}{40}$ in the foetal

* For tables 1 to 6 see pull-out facing p. 240.

sera of the first three of these were to be expected, on the basis of comparable experiments with immune rabbit sera reported previously (Brambell *et al.* 1949), if bovine antibodies behaved like rabbit antibodies. Clearly bovine agglutinins do not pass into the foetal blood by this route as freely as rabbit agglutinins, and no passage could be detected at titres $\frac{1}{4}$ of those expected, which was the limit of sensitivity of the technique. The results did not provide any indication as to whether the differential selection of the bovine as compared to the rabbit agglutinins was effected by the maternal uterine tissues or by the foetal splanchnopleur, nor were the experiments designed to do so.

Intra-uterine injection of rabbit and bovine anti-brucella sera at 24 days post-coitum

Bovine anti-brucella serum was injected into one uterine horn and rabbit anti-brucella serum was injected into the other uterine horn of six rabbits (Nos. 152, 132, 348, 259, 280 and 195) at about 24 days post-coitum. Bovine anti-brucella serum was injected into both uterine horns of one animal (No. 446), and into one uterine horn of another animal (No. 443) in which the other horn was used for other purposes. The injections were performed at laparotomy, the maternal blood being sampled first to ensure that there was no natural immunity to brucellosis. The animals were killed 24 or 48 hr. after injection and both the maternal and foetal sera were titrated. The maternal sera were negative in all cases.

The data are summarized in table 2, the embryos in each uterine horn being numbered from the tubal extremity, where the injection was made, towards the cervix. The sera of all fourteen embryos exposed to immune rabbit sera had positive agglutinin titres of from $\frac{1}{80}$ to $\frac{1}{320}$, whereas of the thirty-two embryos exposed to immune bovine sera, twenty-five were negative at titres of $\frac{1}{10}$ and the titres of none of the seven positives exceeded $\frac{1}{20}$. Since the titre of the immune bovine serum employed was at least double that of the immune rabbit serum, these results are very significant. The embryos which had positive titres were confined to four of the eight litters exposed to immune bovine serum, but embryos which were negative were present in all the uteri. None of the embryos adjoining the sites of injection were positive. It is evident that none of the embryos tested at 24 days of age admitted bovine agglutinins as freely as rabbit agglutinins. Some of the embryos do admit bovine agglutinins in small but significant amounts, and there is a significant individual variation amongst the embryos in this respect. It is important to emphasize, however, that although the majority of the embryos exposed to bovine immune serum were negative at titres of $\frac{1}{10}$ there is no evidence that a more sensitive test would not have revealed the passage of very small quantities of agglutinins. The individual variation of the embryos can be interpreted either as complete exclusion of bovine agglutinins by some embryos, or as a quantitative difference, some admitting significant, and some insignificant, amounts relative to the sensitivity of the test employed.

Injection of rabbit anti-sheep serum

It is believed that the bacterio-agglutinins of rabbits have particles of a molecular weight of approximately 180,000 (ultracentrifugal sedimentation constant of $S_{20} = 7 \cdot 1 \times 10^{-13}$), whereas the bacterio-agglutinins of cattle and horses have much

larger particles of a molecular weight of approximately 930,000 (ultracentrifugal sedimentation constant of $S_{20} = 18 \cdot 0 \times 10^{-13}$). Hence the possibility that the differential selection of rabbit and bovine anti-brucella agglutinins depended on particle size could not be ignored. It was decided to test this possibility by using rabbit antibodies of large particle size, but of the same order as that of the bovine bacterio-agglutinins. Paic (1938) has shown that the haemolysins of rabbits produced in response to injections of the red blood corpuscles of sheep have a sedimentation constant of $18 \cdot 9 \times 10^{-13}$ and hence belong to the group of antibodies of large particle size. Sera of suitably high titre can be obtained.

Rabbit anti-sheep serum was administered intravenously to one rabbit 24 days post-coitum, which was killed after 22 hr. and the sera of the embryos were tested for haemolysins. The results are summarized in table 3. All the embryos were positive, with titres varying from $\frac{1}{8}$ to $\frac{1}{4}$ of that of the maternal serum at the conclusion of the experiment. Hence in this experiment rabbit haemolysins passed from the maternal into the foetal circulations as freely as rabbit agglutinins have been shown to pass (Brambell *et al.* 1949), despite the difference in particle size. The result was so clear that it was decided to seek confirmation by direct comparison of rabbit haemolysin and bovine agglutinin injected respectively into the two uterine horns, rather than by repetition of the intravenous injection experiment. Accordingly, two rabbits were injected, the results being summarized in table 4. The sera of all the embryos exposed to the haemolytic rabbit sera were positive. Although the embryos in one animal were only exposed to the serum for 3 hr., the titres of their sera were surprisingly high, indicating that the mean rate of passage during the first 3 hr. is probably considerably higher than the mean rate of passage over the first 24 hr. The sera of all the embryos exposed to the immune bovine sera were negative.

It was concluded that particle size could not account for the exclusion of bovine agglutinins from the foetal circulation by the yolk-sac splanchnopleur. It appeared possible, however, that the bovine serum, or some component of it, might be having an effect on the foetal yolk-sac splanchnopleur exposed to it which rendered the membrane impervious to globulins. For example, such an effect could be visualized as a blocking of pores in a porous membrane or as a toxic effect on actively absorptive cells. On any such hypothesis globulins presented simultaneously to the yolk-sac splanchnopleur in a mixed serum should be equally acceptable, irrespective of their origin. Intra-uterine injections of mixed rabbit anti-sheep and bovine anti-brucella sera were performed to test this possibility. The results are summarized in table 5. Rabbit haemolysins were present in the foetal sera, often at a high titre, of all the thirty embryos, whereas the bovine agglutinins were not detectable at a titre of $\frac{1}{10}$ in the sera of twenty-nine of these. The serum of one embryo only contained bovine agglutinins at a titre of $+++\frac{1}{10}$, compared to rabbit haemolysins at a titre of $+\frac{1}{40}$, despite the fact that in the intra-uterine serum to which it was exposed, the agglutinin titre was double the haemolysin titre and that this was reinforced by intravenous injection of bovine serum giving an immediate titre of $++\frac{1}{1280}$ and a final titre of $+++\frac{1}{80}$ in the maternal serum. It must be concluded that the yolk-sac splanchnopleur is capable of differentiating rabbit from bovine antibodies

even when these are presented in a common substrate of mixed sera. The yolk-sac splanchnopleur simultaneously can admit rabbit antibodies to, and exclude bovine antibodies from, the foetal circulation. It is interesting that a single embryo in one litter, which displayed no exceptional capacity to admit rabbit haemolysins, should have admitted bovine agglutinins, though relatively much less freely.

Injection of equine anti-diphtheria serum

It was desired to check the conclusion that the passage of antibodies is independent of particle size, by using also foreign antibodies of small particle size, comparable to rabbit agglutinins. Kekwick & Record (1940) have shown that the diphtheria antitoxins of horses are associated with both the β- and γ-globulins with sedimentation constants of $S_{20} = 7 \cdot 18 \times 10^{-13}$ and $S_{20} = 6 \cdot 87 \times 10^{-13}$ respectively, thus belonging to the group of antibodies of small molecular size of about 180,000. They found no antibody reactivity associated with the small quantity of the high molecular component of normal horse serum having a sedimentation constant of $S_{20} = 18 \times 10^{-13}$. Equine serum of high diphtheria antitoxin content can be prepared and the intracutaneous test of antitoxin titre provides a very sensitive and accurate method of assay. Further, Hartley (1948) had demonstrated that diphtheria antitoxin, whether of guinea-pig or horse origin, when administered parenterally to guinea-pigs in late pregnancy is found in the serum of the newborn young. He found that the antitoxin of natural serum was transmitted more readily than that of pepsinized or enzyme-refined serum preparations. Hence equine diphtheria antitoxin was selected as suitable for the following experiments, using only natural untreated serum. The results are summarized in table 6. The equine serum was injected into one uterine horn and rabbit anti-brucella serum into the other horn of four animals (nos. 258, 215, 340, 267). Equine serum was injected into both uterine horns, followed after 24 hr. by an injection of rabbit anti-brucella serum into one horn and bovine anti-brucella serum into the other horn, in one animal (no. 440). Mixed equine and rabbit anti-brucella sera were injected into one horn of the remaining three animals (nos. 441, 311, 500) and mixed equine and bovine anti-brucella sera into the other horn of two of these three (nos. 441, 311). It will be observed that antitoxin was present in significant amounts in the sera of all the embryos exposed to the equine sera. However, the concentrations of antitoxin in the foetal sera were very variable. Expressing the concentrations of the foetal sera as fractions of the concentrations of the injected sera, and comparing the ratios of titre of foetal serum to titre of injected serum for the bovine agglutinins, some allowance can be made for the differences in sensitivity of the methods and an approximate basis obtained for comparison of the permeability of the splanchnopleur to equine antitoxin and bovine agglutinin. The results are summarized on this basis in table 7. Over the range of comparable ratios there is reasonable agreement between the proportions of embryos admitting bovine and equine antibodies respectively.

The results with the mixed sera show no appreciable effect of equine serum on the entry of rabbit or bovine antibodies, neither retarding the one nor accelerating the other. It may be concluded that the results obtained with these experiments were

not due to a toxic action of the equine serum on the foetal membranes. The possibility that the sample of equine serum employed in the experiments on animals nos. 258, 215, 340 and 267 was toxic is not entirely excluded as it was different from that used with the remainder.

TABLE 7. COMPARISON OF ENTRY OF BOVINE AGGLUTININS AND
EQUINE ANTITOXINS INTO FOETAL SERA

ratio: $\dfrac{\text{titre or concentration of foetal serum}}{\text{titre or concentration of injected serum}}$	bovine agglutinins			equine antitoxins		
	$+$	$-$	$+\%$	$+$	$-$	$+\%$
$\frac{1}{64}$	0	68	0·0	0	35	0·0
$\frac{1}{128}$	5	55	8·3	0	35	0·0
$\frac{1}{256}$	9	35	20·5	5	30	14·3
$\frac{1}{512}$	9	21	30·0	13	22	37·1

*Intra-uterine injection of rabbit and bovine anti-brucella sera
at 20 and 22 days post-coitum*

It has been indicated that the individual variation amongst embryos in the entry of bovine agglutinins into the foetal circulation at 24 days could be interpreted as quantitative rather than qualitative. It is reasonable to assume that the capacity to differentiate between homologous and heterologous antibodies appears at a definite stage of development of the embryo. This assumption is supported by the observation that the bilaminar omphalopleur of the 8-day blastocyst admits both rabbit and bovine globulins with equal freedom. It was desired to test the capacity of the yolk-sac splanchnopleur to admit bovine agglutinins at earlier stages of development than 24 days so as to elucidate this problem. Unfortunately, it is not practicable to collect sufficient serum (0·2 ml.) for agglutination titrations from each embryo at ages of less than 24 days. Hence, although serum could be injected into the uterine lumen at any stage, the experiments had to be confined in practice to stages at which the embryos could absorb sufficient quantities of serum to produce a titre in the blood that would still be significant at 24 days. Experiments were performed at 22 and at 20 days, the animals in all cases being killed and the foetal samples taken at 24 days, i.e. 48 or 96 hr. after injection. The results are summarized in tables 8 and 9 respectively. It will be seen that in both series the proportion of embryos on the bovine sides with positive titres is higher than in the 24-day series. The titres of rabbit antibodies in the foetal sera of the embryos on the control sides of the 20-day series vary from negative at $\frac{1}{10}$ to positive at $\frac{1}{80}$, providing only the minimum necessary margin for detection. Hence it was not considered to be profitable to proceed further to still earlier stages.

All the results from experiments on injection into the uterus of either rabbit or bovine anti-brucella sera recorded herein or previously (Brambell *et al.* 1949) are summarized in table 10 according to age at injection and time of exposure. This indicates that the capacity to admit bovine agglutinins tends to decline with increasing age at time of injection both as regards the proportion of embryos with sera giving positive reactions and the maximum titres obtained. Testing the inde-

TABLE 8. INTRA-UTERINE INJECTION OF RABBIT AND BOVINE ANTI-BRUCELLA SERA AT 22 DAYS POST-COITUM

no.	age at operation (days)	(hr.)	time of exposure (hr.)	dose of rabbit serum (ml.)	dose of bovine serum (ml.)	rabbit antibodies embryo 1	titres of bovine antibodies in embryonic sera embryo 1	embryo 2	embryo 3	embryo 4	embryo 5
303	21	19	48	4	7	$+++\frac{1}{160}$	$-\frac{1}{10}$	$+++\frac{1}{20}$	$-\frac{1}{10}$	$++\frac{1}{40}$	$+++\frac{1}{40}$
339	21	20	48	5	7	.	$+++\frac{1}{20}$	$+\frac{1}{80}$	$+++\frac{1}{20}$	$++\frac{1}{40}$	.

TABLE 9. INTRA-UTERINE INJECTION OF RABBIT AND BOVINE ANTI-BRUCELLA SERA AT 20 DAYS POST-COITUM

no.	age at operation (days)	(hr.)	time of exposure (hr.)	dose of rabbit serum (ml.)	(titre)	dose of bovine serum (ml.)	(titre)	titres of rabbit antibodies in embryonic sera embryo 1	embryo 2	embryo 3	titres of bovine antibodies in embryonic sera embryo 1	embryo 2	embryo 3	embryo 4	embryo 5	embryo 6
281	20	.	96	7	$\frac{1}{2560}$	7	$\frac{1}{5120}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	$+\frac{1}{40}$	$+++\frac{1}{10}$	$+\frac{1}{40}$	$++\frac{1}{20}$		
278	20	0	98	7	$\frac{1}{2560}$	7	$\frac{1}{5120}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	$+\frac{1}{80}$	$-\frac{1}{10}$	$\frac{1}{20}$	$-\frac{1}{10}$	$-\frac{1}{10}$	
243	20	0	96	8	$\frac{1}{2560}$	6	$\frac{1}{2560}$	$+++\frac{1}{80}$	$+++\frac{1}{80}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.
331	19	21	94	7	$\frac{1}{1280}$	7	$\frac{1}{2560}$	$+++\frac{1}{40}$	$+++\frac{1}{40}$	$+\frac{1}{80}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	$-\frac{1}{10}$	.
387	19	17	97	5	$\frac{1}{1280}$	7	$\frac{1}{2560}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	.	$-\frac{1}{10}$	$-\frac{1}{10}$	$+\frac{1}{40}$	$-\frac{1}{40}$	$-\frac{1}{10}$	$-\frac{1}{10}$
245	19	16	96	7	$\frac{1}{1280}$	7	$\frac{1}{2560}$	$+++\frac{1}{20}$	$+++\frac{1}{20}$	$+\frac{1}{20}$	$-\frac{1}{10}$	$+\frac{1}{20}$	$+\frac{1}{40}$	$+\frac{1}{40}$	$-\frac{1}{10}$	$+\frac{1}{40}$

pendence of the proportions of positive and negative results from injection of bovine serum at 20 days and at 24 days as a 2×2 classification, $\chi^2 = 9 \cdot 97$ and $P < 0 \cdot 01$. Hence the proportion of positives resulting from injection at 24 days is very significantly less than the proportion resulting from injection at 20 days. Conversely, the capacity to admit rabbit antibodies appears to increase with increasing age from 20 to 24 days as judged by the minimum titres observed. This increase may be apparent rather than real. If, as seems probable, absorption takes place in the first few hours after injection, before the injected serum has had time to drain away through the cervix, the amount of subsequent dilution of the foetal serum, resulting in corresponding diminution of the final titre would be proportional to the lapse of time. This consideration does not detract from the significance of the result that whereas in the 20-day series the observed range of titres of bovine and rabbit agglutinins are coincident, at 24 days they barely overlap. This decline in the admission of bovine agglutinins as compared to rabbit agglutinins is more significant than the absolute decline. It strongly supports the hypothesis that the embryos are gradually developing from approximately the 20th day onwards the capacity to exclude bovine globulins and that, in consequence, owing to normal developmental variation of the embryos a declining proportion give positive reactions at a serum dilution of $\frac{1}{10}$.

TABLE 10. SUMMARY OF INTRA-UTERINE INJECTIONS OF RABBIT OR
BOVINE ANTI-BRUCELLA SERA ACCORDING TO AGE

age at operation (days)	time of exposure (hr.)	serum used	no. of litters	no. of embryos observed with sera of given titre						
				$-\frac{1}{10}$	$+\frac{1}{10}$	$+\frac{1}{20}$	$+\frac{1}{40}$	$+\frac{1}{80}$	$+\frac{1}{160}$	$+\frac{1}{320}$
20	96	rabbit	6	2	1	10	1	3	·	·
		bovine	6	15	1	3	6	1	·	·
22	48	rabbit	1	·	·	·	·	·	1	·
		bovine	3	5	·	3	2	1	·	·
24	24	rabbit	18	·	·	4	15	6	16	4
		bovine	15	55	4	5	·	·	·	·
24	48	rabbit	3	·	·	·	·	1	5	4
		bovine	2	6	·	·	·	·	·	·
26	24	rabbit	1	·	·	·	1	·	1	·
		bovine	·	·	·	·	·	·	·	·
26	48	rabbit	2	·	·	·	·	2	1	1
		bovine	·	·	·	·	·	·	·	·

TITRE OF FOETAL SERUM RESULTING FROM INTRA-UTERINE
INJECTION OF IMMUNE RABBIT SERUM

It is evident that the antibody titre in the foetal sera is closely related to that of the injected serum in the experiments on intra-uterine injection of rabbit antibodies. The most extensive data are those for intra-uterine injection at 24 days, derived partly from experiments described previously (Brambell *et al.* 1949) and partly from those described herein. These data, including both those for the injection of anti-brucella and anti-sheep rabbit sera, are summarized in table 11. It can be seen by inspection that there is a close correlation between the titre of the injected serum

and that of foetal sera. This is brought out clearly if the titres of the foetal sera are expressed as fractions of the titres of the injected sera, as in table 12. It is apparent that the range of variation falls within two dilutions on either side of the maximum frequency of 1/16th. Corresponding data for the experiments on intravenous injection of immune rabbit sera at 24 days (Brambell et al. 1949) are included for comparison, the titres of the foetal sera being expressed as fractions of the titres of the maternal sera at the times of killing. The titre of the maternal serum falls, from a maximum immediately after injection to a minimum at the conclusion of the experiment, so that the average titre of the maternal serum during the experiment would tend to be slightly higher, and the fractional ratio of the foetal serum correspondingly smaller. Moreover, whereas immune serum injected into the uterine lumen probably drains away gradually through the cervix, so that the foetal membranes may be exposed to it for only a part of the period of the experiment, they will be exposed throughout the whole period of the experiment to maternal circulating antibodies. These two considerations probably account for the apparent greater efficiency of absorption of antibodies from the maternal circulation than from the uterine lumen. The regularity of entry of rabbit antibodies into the foetal circulation is apparent in both series of data.

TABLE 11. TITRES OF FOETAL SERA AFTER INTRA-UTERINE INJECTION
OF IMMUNE RABBIT SERUM AT 24 DAYS POST-COITUM

titre of serum injected	titre of foetal serum						
	$\frac{1}{10}$	$\frac{1}{20}$	$\frac{1}{40}$	$\frac{1}{80}$	$\frac{1}{160}$	$\frac{1}{320}$	$\frac{1}{640}$
$\frac{1}{20480}$	.	.	.	.	.	4	.
$\frac{1}{10240}$	.	.	.	.	.	.	.
$\frac{1}{5120}$	.	.	.	.	2	1	2
$\frac{1}{2560}$	.	.	1	4	12	9	.
$\frac{1}{1280}$	.	1	8	8	5	.	.
$\frac{1}{640}$	1	2	4	.	5	.	.
$\frac{1}{320}$	.	5	11	1	.	.	.

TABLE 12. TITRES OF FOETAL SERA EXPRESSED AS FRACTIONS OF TITRES OF
IMMUNE SERA TO WHICH THEY WERE EXPOSED AT 24 DAYS POST-COITUM

ratio: $\dfrac{\text{titre of foetal serum}}{\text{titre of maternal or intra-uterine serum}}$	intra-uterine injection experiments	intra-venous injection experiments
$\frac{1}{64}$	7	.
$\frac{1}{32}$	16	.
$\frac{1}{16}$	30	1
$\frac{1}{8}$	27	8
$\frac{1}{4}$	6	13
$\frac{1}{2}$	.	5

CONCLUSIONS

The yolk-sac splanchnopleur of foetal rabbits at 24 days admits rabbit antibodies freely to the foetal circulation but admits bovine or equine antibodies relatively slightly or not at all. This selection cannot be a simple sieving effect governed by

particle size, since rabbit antibodies of widely divergent molecular weights are admitted equally freely whereas equine and bovine antibodies of corresponding molecular dimensions are partially or totally excluded. Hartley's (1948) finding that in guinea-pigs pepsin-refined diphtheria antitoxins, whether of guinea-pig or equine origin, when injected parenterally into pregnant females did not pass into the foetal circulation, whereas the antitoxins of the natural sera did, supports the hypothesis, since reduction in size of the particles was accompanied by decrease in the rate of passage.

Moreover, when the yolk-sac splanchnopleur is exposed to mixtures of the two sera, there is no evidence of any mutual interference causing increase or decrease in the admission of rabbit, bovine or equine antibodies. It appears that the selection depends on molecular characters, other than mass, of the γ-globulins themselves.

The selectivity of the yolk-sac splanchnopleur is developing progressively between the 20th and 24th days. During this period the capacity to admit bovine globulins, measured in terms of the capacity to admit rabbit globulins, declines. The stage at which this immunological character is developed is well defined and is relatively early in ontogeny.

It must be realized that all the evidence points to a decline in the capacity of the splanchnopleur to admit heterologous globulins to a level below that at which the sensitivity of the methods employed permits detection. Thus although only a small proportion of rabbit embryos exposed to immune bovine serum at 24 days contain bovine antibodies in their sera in sufficient concentrations to be detectable by agglutination tests at dilutions of $\frac{1}{10}$, it may be that all contain some. The intra-cutaneous test for diphtheria antitoxin is capable of determining much lower concentrations of the antitoxin. Taking this into account the finding of diphtheria antitoxin in the sera of all the embryos exposed at this stage of development to equine anti-diphtheria immune sera, does not conflict with the bovine results. It should be remembered also, in this connexion, that Winter (1944) has shown that even in the adult guinea-pig sufficient equine serum proteins can be absorbed from the small intestine to desensitize animals previously sensitized to these proteins.

The significant individual variation between embryos even in the same litter, in their capacity to admit bovine agglutinins, is interesting. It suggests that this is a genetic character and hence supports the view that selectivity is a cellular phenomenon.

The selectivity of the splanchnopleur appears to be in line with the specificity characteristic of other serological reactions. However, the production, after an interval of some days, of antibodies to specific antigens, does not appear comparable since the selectivity of the foetal splanchnopleur is displayed immediately on first exposure to foreign antibodies. It resembles more closely the phenomenon of passive anaphylactic sensitization in guinea-pigs, which follows within a few hours after the intravenous injection of immune guinea-pig or rabbit serum. Administration of the antigen to the sensitized animal is followed immediately by the well-known and characteristic symptoms of anaphylaxis. Although Bailey, Raffel & Dingle (1937) claim to have passively sensitized guinea-pigs with equine

serum, many other workers have failed to do so with equine and other ungulate sera. A particularly sensitive and convenient means of demonstrating sensitization in guinea-pigs is provided by isolated smooth muscle preparations. Dale (1913) has shown that the smooth muscle of the isolated uterus of sensitized virgin guinea-pigs responds immediately by maximal contraction to the introduction into the surrounding saline medium of minute traces of the homologous antigen but not of other antigens. Hartley (1939), using Dale's technique, showed that the uterine smooth muscle of normal guinea-pigs could be passively sensitized *in vitro* by suspension in oxygenated Ringer solution containing suitable antibodies. Diphtheria antitoxin from guinea-pig, rabbit or man sensitized the uteri, whereas diphtheria antitoxin from the horse, ox, goat, sheep and pig did not. He found (1948) that a very few units of diphtheria antitoxin prepared in a guinea-pig were sufficient to produce sensitization, whereas many thousands of units of equine antitoxin failed to do so. Hartley compared these results to his finding that guinea-pig antitoxin passed more readily than equine antitoxin from the maternal to the foetal circulation in guinea-pigs.

The phenomenon of sensitization is generally interpreted as being due to the fixation of previously circulating antibodies on or in the cells. It can persist even after the antibody has disappeared from the circulation. Fixation in this sense can be defined as the uptake by the cells from the surrounding medium of antibodies and their retention, as such, on or in the cells. Anaphylaxis is believed to be brought about by the reaction of the antigen with these fixed antibodies, damaging the cells and resulting in the release from them of histamine and possibly other toxic substances. Essentially it is an antibody-antigen reaction on or in the cells, as distinct from such reaction in the circulation. Indeed large quantities of circulating antibody will protect a sensitized animal from anaphylaxis on intravenous injection of doses of antigen that otherwise would kill it, by neutralizing the antigen in the blood. That the cells are, nevertheless, still sensitive in such animals can be shown by isolated smooth muscle preparations. Thus we arrive at the conception that guinea-pig smooth muscle cells can fix, in the sense defined, antibodies of guinea-pig or rabbit origin but not those of ungulate origin. Whereas the rodent antibodies are retained in the cell in a reactive form, the ungulate antibodies are either excluded or permitted to enter but not retained in a reactive form. If these ungulate antibodies do enter, then they must either emerge again before an anaphylactively reactive concentration is attained or else be inactivated, presumably by being metabolized. This conception of sensitization can be represented diagramatically as in figure 1 a. It is difficult to avoid the assumption that the selectivity of the yolk-sac splanchnopleur is also a cellular phenomenon. The endodermal epithelium and the mesodermal endothelium of the foetal blood vessels are the only histologically continuous membranes in the yolk-sac splanchnopleur intervening between the uterine lumen and the foetal circulation, hence one or other or both of these must constitute the barrier. We know that rabbit antibodies are permitted to penetrate this barrier freely while retaining their reactivity, whereas bovine and equine antibodies are relatively almost excluded. This exclusion must be at the surface of the barrier or the antibodies are permitted to enter the substance of the barrier without emergence

in a reactive form, being either retained or inactivated. This is represented diagrammatically in figure 1b. The phenomena both of anaphylactic sensitization and of membrane permeability could be accounted for by assuming that the foreign antibodies are either excluded from the cells or are rapidly inactivated upon entering them, whereas the homologous antibodies can be both fixed and subsequent released without inactivation. This hypothesis is represented diagrammatically in figure 1c. It has the merit that it would accord well with the views of Whipple (1948) based on physiological rather than immunological evidence, that the plasma proteins pass readily into and out of the cells and that a state of dynamic equilibrium exists between cell and plasma protein.
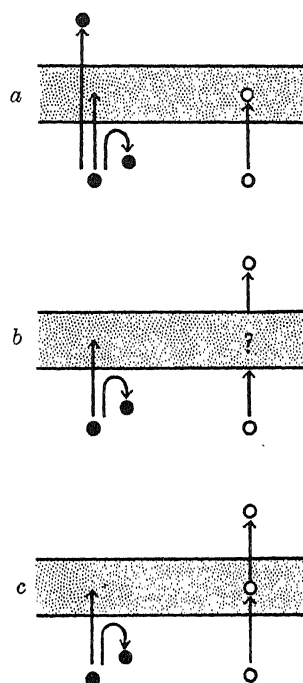
FIGURE 1. *a*, diagrammatic representation of anaphylactic sensitization of guinea-pig smooth muscle. *b*, passage of antibodies through the foetal yolk-sac splanchnopleur of rabbits. *c*, an interpretation of both sets of phenomena. Homologous antibody molecules are represented by hollow circles, heterologous antibody molecules by solid circles and the cells by the stippled area.

REFERENCES

Bailey, G. H., Raffel, S. & Dingle, J. H. 1937 *Amer. J. Hyg.* **25**, 381.
Brambell, F. W. R., Hemmings, W. A. & Rowlands, W. T. 1948 *Proc. Roy. Soc.* B, **135**, 390–403.
Brambell, F. W. R., Hemmings, W. A., Henderson, M., Parry, H. J. & Rowlands, W. T. 1949 *Proc. Roy. Soc.* B, **136**, 131–44.
Dale, H. H. 1913 *J. Pharmacol.* **4**, 167–223.
Hartley, P. 1939 *Int. Congr. Microbiol.* **3**, 343. New York.
Hartley, P. 1948 *Mon. Bull. Minist. Hlth*, **7**, 45–53.
Kekwick, R. A. & Record, B. R. 1940 *Brit. J. Exp. Path.* **22**, 29–44.
Paic, M. 1938 *C.R. Acad. Sci., Paris*, **207**, 1074–6.
Whipple, G. H. 1948 *Haemoglobin plasma protein and cell protein.* Springfield, Illinois, U.S.A. Charles C. Thomas.
Winter, L. B. 1944 *J. Physiol.* **102**, 373–405.

The phosphorus metabolism of the brain

By R. M. C. Dawson, *Beit Memorial Fellow* and D. Richter

The Neuropsychiatric Research Centre, Whitchurch Hospital, Cardiff

The metabolism of the brain *in vivo* has been studied by measuring the rate of uptake of radioactive phosphorus into the different phosphorus-containing fractions of the mouse brain. By the use of specific-activity ratios referred to the acid-soluble fraction of the brain, satisfactorily constant values were obtained for the uptake into the nucleoprotein and phospholipid fractions in normal animals. The observed ratios indicated a relatively high metabolic activity for these fractions under normal conditions *in vivo*.

The uptake of radioactive phosphorus into the nucleoprotein and phospholipid fractions of the brain was decreased in sodium pentobarbital anaesthesia; the effect was greater if the body temperature was also allowed to fall. Electrically induced convulsions and insulin hypoglycaemia caused a significant decrease in the uptake of radioactive phosphorus into the brain phospholipids without a corresponding change in the nucleoprotein fraction. A similar specific effect on the phospholipid metabolism was observed under normal physiological conditions in animals exposed for 3 hr. in a rotating drum. The effect was absent in animals which had previously been conditioned to the rotating drum. The results give evidence that the metabolism of the permanent or 'structural' elements of nervous tissue, as well as of the more labile metabolites, may vary *in vivo* with the state of functional activity of the brain.

INTRODUCTION

There is reason to believe that the metabolic activity of the brain *in vivo* is not constant under all conditions, but varies with the state of functional activity of the brain. The rate of utilization of oxygen and glucose, for example, as indicated by the arterio-venous differences in the composition of the blood entering and leaving the brain, is decreased in anaesthesia and increased in convulsions (Dameshek, Myerson & Loman 1934; Gurdjian, Webster & Stone 1947). Direct determinations

of the metabolites in the brain, after rapid fixation of the tissue by freezing with liquid air, show an increase in the acetylcholine content and a fall in the lactic acid in anaesthesia and in sleep, while the converse is true in states of increased functional activity of the brain (Richter & Dawson 1948; Richter & Crossland 1949). The phosphate esters also show significant changes in different states of functional activity; the rapid breakdown of phosphocreatine and increase in the 'hexose phosphate' fraction of the brain in convulsions may be associated with the increase in the carbohydrate utilization due to increased activity of the energy-supplying mechanisms (Dawson & Richter 1949).

In the present investigation the phosphorus metabolism was studied by following the pathway taken by the radioactive isotope, ^{32}P, as it enters the brain from the blood stream and becomes incorporated into the different phosphorus-containing fractions of the brain. In this way the dependence of the brain metabolism *in vivo* on the state of functional activity was confirmed by an independent method. It was further shown that the metabolic changes with functional activity *in vivo* are not confined to the labile metabolites such as the phosphate esters, but they also involve the more stable or 'structural' elements of nervous tissue.

THE ENTRY OF RADIOACTIVE PHOSPHORUS INTO THE BRAIN

When radioactive phosphorus in the form of inorganic phosphate is administered to an animal by intraperitoneal or subcutaneous injection, the phosphate rapidly enters the blood stream. Some of the injected phosphate enters the blood corpuscles where it is synthesized into phosphate esters, but the major part passes into the muscles and other tissues of the body. Here the radioactive phosphate (^{32}P) is rapidly exchanged with the normal phosphate (^{31}P) and enters into the general metabolism of the tissues.

During the first few hours after administration the radioactive phosphorus is found mainly in the acid-soluble phosphate ester fraction; but at the same time it participates in the much slower synthesis of nucleoprotein and phospholipid. The initial rise in the specific activity ($^{32}P/^{31}P$) in the tissues is followed by a steady fall as radioactive phosphorus is lost by excretion and by slow migration into the inorganic deposits of the skeleton.

The rate of entry of radioactive phosphorus into the brain is relatively slow. Hevesy (1939) found only 0·02 % of the injected phosphorus in the brain after 4 hr., and Cohn & Greenberg (1938) found that the brain of the rat gave the least retention of the isotope per unit weight of any organ. These observations have led to the view that the phosphorus metabolism of the brain is slower than that of other tissues; but it has recently been shown that radioactive phosphorus is quickly synthesized into phosphate esters in the brain (Lindberg & Ernster 1950), and the present work gives evidence that the turnover rate in the nucleoprotein and phospholipid fractions of the phosphorus which enters the brain is in fact relatively rapid. The slow entry of radioactive phosphorus into the brain cannot therefore be attributed to metabolic inactivity, and it would appear more likely that it is due to the low permeability of the blood-brain barrier to phosphate ions.

THE METHOD OF DETERMINING THE RATE OF PHOSPHORUS TURNOVER

Young adult mice were used as experimental animals in the present work. After injecting the radioactive phosphate, a period of 3 hr. was allowed for phosphorus exchange. The mice were then decapitated; the brains were rapidly removed and ground in a mortar with trichloroacetic acid.

The phosphorus compounds of the brain were separated into three main fractions:

(a) the *acid-soluble fraction*, containing the nucleotides, inorganic phosphate and phosphate esters soluble in 5 % aqueous trichloroacetic acid solution,

(b) the *nucleoprotein fraction*, containing ribo- and desoxyribo-nucleic acids, with a smaller amount of phosphoprotein, of which relatively little is present in the brain (Schmidt & Thannhauser 1945), and

(c) the *phospholipid fraction*, containing lecithins, cephalins and sphingomyelins extracted by alcohol-chloroform mixture and sparingly soluble in acetone.

The specific activity ($^{32}P/^{31}P$) of the phosphorus in these three fractions was determined by measuring the ratio of radioactive phosphorus to the total amount of phosphorus present. The specific activity figures have commonly been used as an indication of the rate of turnover of the phosphorus in any fraction over a given period of time; but the specific activity shows considerable variation from one animal to another, since it depends on the absolute amount of radioactive phosphorus entering the tissue. It is therefore affected by individual differences in such factors as the rate of absorption from the site of injection, the loss by excretion and the rate of entry into the other tissues of the body. For the same reason the specific activity figures cannot be taken as giving a reliable measure of the metabolic activity or of the rate of synthesis of the fraction.

Hevesy & Hahn (1940) have used the ratio of the specific activity of the phosphorus in the fraction to the specific activity of the inorganic phosphate in the blood plasma or in the tissue. The *specific-activity ratio* obtained in this way has the advantage of being independent of the absolute amount of radioactive phosphorus in the tissue; it gives the rate of transfer of radioactive phosphorus into the fraction, and hence it gives a valuable indication of the metabolic activity. It may be noted, however, that in determining the activity ratios of fractions such as the nucleoprotein or phospholipid fractions of the brain there is no theoretical justification for using the specific activity of the inorganic phosphate rather than that of any other component of the acid-soluble fraction as the denominator in the ratio. The immediate precursors of the nucleoprotein and phospholipid fractions are not known, but their phosphorus is as likely to come from the high-energy phosphate esters as from inorganic phosphate, of which very little is present *in vivo* in the brain. There is a relatively rapid exchange of phosphorus between the various components of the acid-soluble fraction, so that they soon come into mutual equilibrium. The specific activity of the whole acid-soluble fraction was therefore used in the present work as the denominator in determining the specific-activity ratios. This made unnecessary the separate determination of the inorganic phosphate of the brain. Three specific-activity ratios were determined:

(a) $\dfrac{\text{Specific activity of brain acid-soluble phosphorus}}{\text{Specific activity of blood acid-soluble phosphorus}}$. This activity ratio indi-

cates the rate of entry of phosphate ion from the blood into the brain during the experimental period.

(b) $\dfrac{\text{Specific activity of brain nucleoprotein phosphorus}}{\text{Specific activity of brain acid-soluble phosphorus}}$. This gives the rate of transfer of radioactive phosphorus from the acid-soluble fraction into the nucleoprotein fraction: hence it indicates the rate of phosphorus turnover of the nucleoprotein fraction during the experimental period.

(c) $\dfrac{\text{Specific activity of brain phospholipid phosphorus}}{\text{Specific activity of brain acid-soluble phosphorus}}$. This ratio indicates the rate of phosphorus exchange of the phospholipid fraction. The three ratios are largely independent of the actual amount of radioactive phosphate entering the brain.

EXPERIMENTAL PROCEDURE

Adult mice 3 to 6 months old and weighing 25 to 38 g. were used in the present work; littermates matched for sex and weight were used as far as possible in the comparison of experimental animals with controls. The radioactive phosphorus solution was given by subcutaneous or intraperitoneal injection, using an all-glass micrometer syringe with a lead cylinder round the barrel to protect the hands. The radioactive phosphorus was obtained from Oak Ridge, U.S.A., or Harwell, England, as a solution of sodium phosphate or phosphoric acid in a glass ampoule. This was centrifuged to remove solution from the neck, and the solution was transferred by means of a teat pipette to a graduated centrifuge tube cut off at the 3 ml. mark. The solution for injection was made by diluting to a final concentration of $50\,\mu$curies/ml. with physiological saline containing sodium dihydrogen phosphate (0.1 mg. P/ml.) as a carrier. The solution injected (0.2 ml.) contained 0.02 mg. P in the form of phosphate and $10\,\mu$curies of radioactive phosphorus for each mouse. It was shown by Jones (1948) that the proportion of injected radioactive phosphorus recovered from the mouse brain is constant in doses ranging from 1 to $70\,\mu$curies; Tuttle, Erf & Lawrence (1941), who investigated the maximum tracer dose permissible without interference with the phosphorus metabolism in the mouse, concluded that $10\,\mu$curies is a safe dose.

After the 3 hr. allowed for the exchange of phosphorus, the mouse was decapitated. About 0.15 ml. of blood from the carotid arteries was stirred into 15 ml. of 5 % trichloroacetic acid in a weighed tube at 0° C. After shaking and standing for 15 min. at 0° C the mixture was filtered; in this way a solution of the acid-soluble phosphorus compounds of the blood was prepared. Meanwhile, the whole brain was quickly dissected from the decapitated head. The brain was gently turned over on dry filter paper to allow as much as possible of the blood to drain from the tissue; it was then quickly weighed in a torsion balance and thoroughly ground in a mortar with 10 ml. of 5 % trichloroacetic acid. The whole procedure took $1\frac{1}{2}$ to 2 min. from the time of decapitation to the grinding with trichloroacetic acid. The suspension was transferred to a 35 ml. capacity centrifuge tube and the residue in the mortar washed in with a further 5 ml. of 5 % trichloroacetic acid. After standing for not less than 15 min. in the refrigerator, the mixture was centrifuged. The supernatant

solution containing the acid-soluble compounds of the brain was filtered and kept for determination of the radioactive and total phosphorus content.

The precipitate containing proteins and lipids was washed once with 30 ml. of 5 % trichloroacetic acid and left standing overnight in the refrigerator with 30 ml. of 1 % trichloroacetic acid. Preliminary tests showed that no more acid-soluble phosphate could be removed by further washing, but the residue was washed twice again with 30 ml. water to remove traces of trichloroacetic acid. The precipitate was then treated with 10 ml. acetone. This removed water and some of the neutral fat and cholesterol. A little phospholipid was also removed by the acetone, probably in the form of lecithin, which dissolves to some extent in acetone solutions of neutral fat (Maclean 1914). Preliminary experiments showed that the specific activity of this acetone-soluble phospholipid did not differ significantly from that of the main phospholipid fraction, so that there was no need to introduce a correction on this account.

The residue left after centrifuging and decanting the acetone solution was dried slowly; it was then extracted by refluxing in the original centrifuge tube with 11 ml. of a 50 % by volume alcohol-chloroform mixture. This removed 90 to 95 % of the phospholipid phosphorus. The solution was decanted through a filter and the filtrate was made up to 10 ml. with the same solvent. The solution was then kept until the radioactivity and the total phosphorus content were determined.

The residual precipitate was washed by refluxing for 30 min. with 20 ml. of a solution containing 75 % by volume of alcohol and 25 % ether. The removal of phospholipids was completed by three extractions for 1 hr. each with 20 ml. of a boiling 50 % by volume alcohol-chloroform mixture. It was found that the final extraction removed only a negligible amount of phosphorus. The residue containing nucleoprotein and phosphoprotein was brought into solution by oxidation with 2 ml. of 70 % perchloric acid. The tube was cautiously heated until solid matter had disappeared, when a few drops of 100 vol. hydrogen peroxide were added to complete the oxidation of coloured organic matter. Goggles were worn and other precautions were taken to avoid injury during the oxidation with perchloric acid, since occasionally the tube exploded violently. After dilution to about 8 ml. the solution was filtered to remove a white precipitate; preliminary tests showed that the radioactive phosphorus content of this precipitate was negligible. The filtrate was made up to 10 ml. with water and kept for analysis; it contained approximately 10 % by volume of perchloric acid.

THE ASSAY OF RADIOACTIVE PHOSPHORUS

Solutions were assayed for radioactivity in a Geiger-Müller counter tube designed for use with liquids (Veall 1948). The tube was contained in a cylindrical lead castle and connexion was made through a mercury cup with the central wire, which was charged to about 80 V above the counting threshold voltage by a power pack (M.R.C. type 200). Pulses from the counter tube were passed into a probe unit (M.R.C. no. 104). The amplified pulses were then passed directly into a scaling unit (M.R.C. type 200) which recorded the number of pulses arriving in a given time. The observed counting rate increased with the volume of the sample up to a certain

'critical volume' of approximately 9 ml. characteristic for the counter tube; there was then no further increase in the counting rate when this volume was exceeded. The sample solution (10 ml.) was introduced, and the counting rate corresponding to the critical volume was then recorded.

The accuracy of the observed counting rate depends on the total number of counts, in accordance with the relation $E_n = 0.67\,N$, where N is the total number of counts recorded and E_n is the probable error. Samples giving low counting rates were therefore observed for a longer period than those giving high counts. The probable error for high counting rates was of the order of 2 % and for the lowest rates not more than 4 %.

The specific activity of a fraction was taken as the number of counts recorded from a 10 ml. sample in the counter tube divided by the number of milligrams of phosphorus contained in the sample. The total phosphorus content was determined after the assay of radioactivity by evaporating an aliquot of the sample nearly to dryness, oxidizing with perchloric acid (King 1946) and estimating the inorganic phosphate by the method of Fiske & Subbarow (1929). Suitable aliquots for this estimation were 2 ml. each of the brain acid-soluble and nucleoprotein phosphorus solutions, 1.5 ml. of the phospholipid solution and 5 ml. of the blood acid-soluble phosphorus solution. Further experimental details and control experiments are described by Dawson & Richter (1949). Throughout the experimental period the usual precautions were taken to avoid radiation hazards.

Corrections

A number of corrections were necessary. The radioactive decay during a series of readings was corrected for by observing the counting rate of a standard solution of radioactive phosphorus every few hours. This also served to correct for any slight changes in the sensitivity of the counter tube. Corrections were made in the usual manner for the background count and for the 'dead time' of the counter tube.

Correction for solvent. The observed counting rate of radioactive phosphorus in solution depended on the degree of absorption of β-particles by the solvent. In order to correct for this, the counting rate of a standard amount of radioactive phosphorus in distilled water was compared with the counting rate of the same amount of radioactive phosphorus in solvents of similar composition to the solutions obtained experimentally in the extraction of the different fractions of the brain. The results are given in table 1. The figures obtained in this way were used to correct the counting rates observed in the metabolism experiments; all counting rates were thus reduced to the rate which would have been obtained if the radioactive phosphorus had been dissolved in an equal volume of distilled water.

Correction for the blood content of the brain. The blood remaining in the brain after decapitation and draining was estimated by the method of Ashby & Chan (1943). Estimations on four brains gave 1.16, 1.32, 0.99 and 1.36, with a mean of 1.21 g. of blood per 100 g. fresh mouse brain.

The specific activities of the phosphorus fractions of blood samples taken 3 hr. after the injection of radioactive phosphorus were determined. The specific activity of the blood nucleoprotein phosphorus was five to eight times that of the brain

nucleoprotein phosphorus, while the blood phospholipid phosphorus had a specific activity approximately twenty times higher than the phospholipid phosphorus of the brain. However, the amounts of blood nucleoprotein and phospholipid in the brain are small in comparison with the relatively large total amounts of these compounds in the brain; approximate calculations showed that the presence of blood nucleoprotein and phospholipid phosphorus of high specific activity would not affect the observed values for the brain nucleoprotein and phospholipid phosphorus activities by more than 2 %. The specific activity of the blood acid-soluble phosphorus after 3 hr. was about ten times that of the brain acid-soluble phosphorus. No correction was applied, but these figures indicate that approximately 5 % of the observed values for the specific activity of the brain acid-soluble phosphorus should be attributed to the blood present in the brain.

TABLE 1. THE COUNTING RATE OF A STANDARD AMOUNT OF
RADIOACTIVE PHOSPHORUS IN DIFFERENT SOLVENTS

solvent	specific gravity	counts per min.	counting rate as % of rate in water
water	1·000	1546	100
5 % trichloroacetic acid	1·025	1515	97·9
10 % perchloric acid	1·096	1432	92·3
ethyl alcohol-chloroform: equal volumes	1·144	1488	96·2
acetone	0·797	1745	113·0

A correction was applied for the small quantity of water introduced into the solvent on adding the standard aliquot of radioactive phosphorus.

THE RATE OF UPTAKE OF RADIOACTIVE PHOSPHORUS IN THE MOUSE BRAIN

The rate at which radioactive phosphorus passes from the blood into the acid-soluble phosphorus compounds of the brain was studied in a series of mice, which were killed at different times after intraperitoneal injection of 10 μcuries of radioactive phosphate in 0·02 mg. of sodium dihydrogen phosphate. In this experiment, and in subsequent work, the mice were examined in series of four at a time; except for animals subjected to insulin hypoglycaemia, they were allowed free access to food and water during the experimental period.

The specific activities of the phosphorus in the acid-soluble fractions of the blood and of the brain were determined; the results, expressed in terms of the specific-activity ratios, are given in figure 1. The rate of increase in the activity ratio of brain to blood was not constant during the experimental period of 6 hr., but showed a falling off after the first hour. This may be due in part to the synthesis of inorganic phosphate into the nucleotides and phosphate esters, which pass less readily than phosphate ions through the blood-brain barrier.

The rates of uptake of radioactive phosphorus into the nucleoprotein and phospholipid fractions in the same series of animals are shown in figures 2 and 3. As was to be expected, the specific activity of the phosphorus in the nucleoprotein and phospholipid fractions was considerably lower than in the acid-soluble fraction of

the brain; but the activity ratios indicate a high metabolic activity for fractions which have generally been regarded as relatively inert or structural elements of the tissue. It can be calculated that during the 3 hr. experimental period about 3·1 mg. P/100 g. fresh brain tissue per hour of the phospholipid phosphorus is
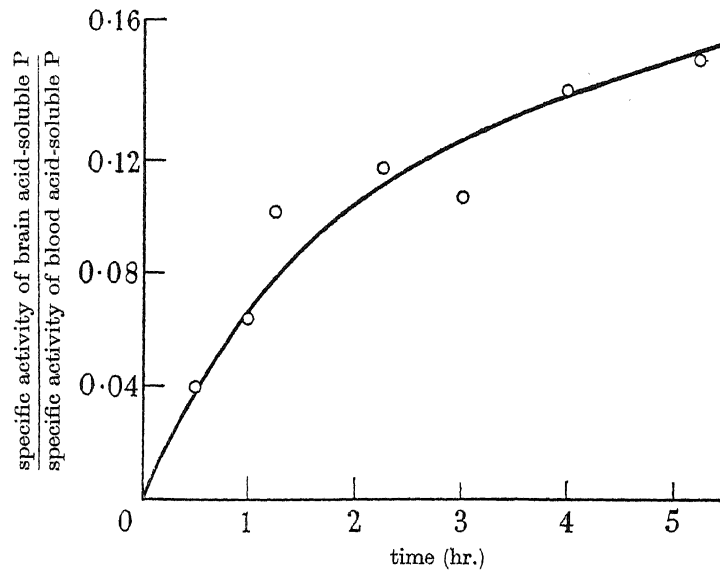
FIGURE 1. The uptake of radioactive phosphorus from the blood into the acid-soluble phosphorus compounds of the mouse brain.
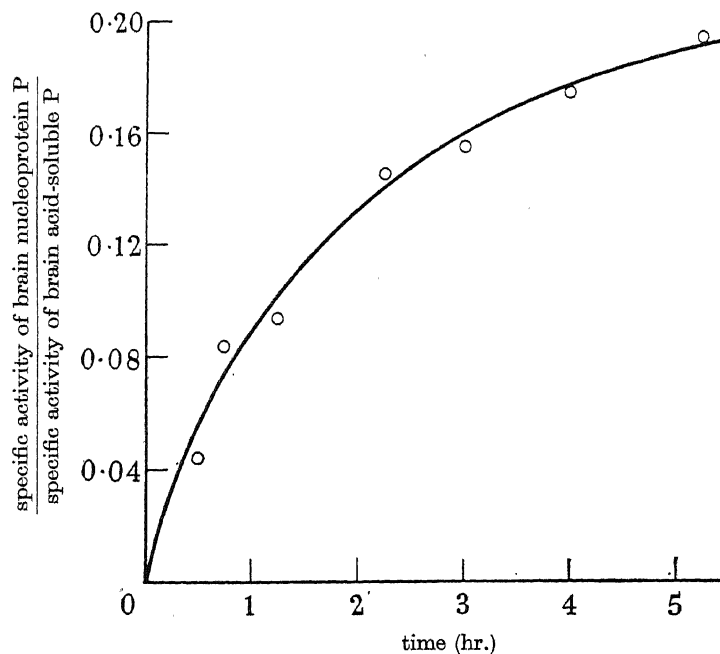
FIGURE 2. The uptake of radioactive phosphorus into the nucleoprotein fraction of the brain.

exchanged with that of the acid-soluble fraction in the brain. This indicates that an amount of phosphorus equivalent to the total phospholipid phosphorus of the brain would be exchanged in approximately 70 hr.
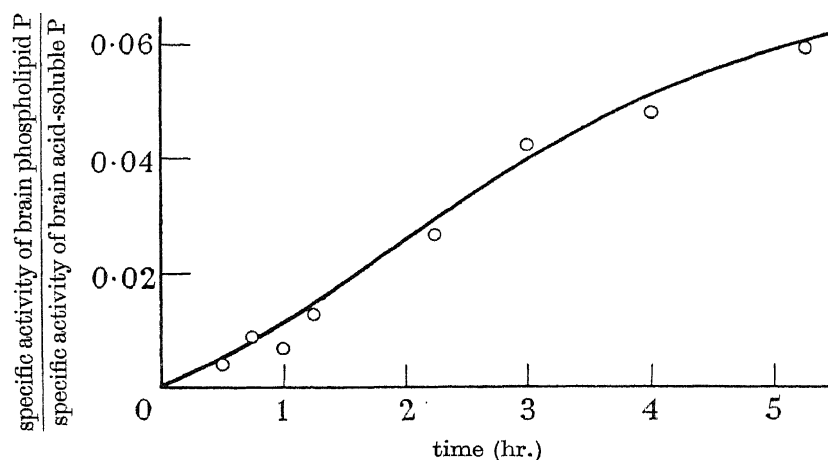
FIGURE 3. The uptake of radioactive phosphorus into the phospholipids of the mouse brain.

The effects of post-mortem autolysis and of age

The previous work suggested the choice of 3 hr. as a suitable experimental period for the comparison of groups of normal animals with animals in different physiological states; but before a satisfactory normal series could be obtained it was necessary to make sure that the figures were not affected by post-mortem changes

TABLE 2. (a) THE EFFECT OF AGE ON THE UPTAKE OF RADIOACTIVE PHOSPHORUS BY THE MOUSE BRAIN, AND (b) THE EFFECT OF POST-MORTEM AUTOLYSIS

mouse no.	weight (g.)	age (days)	time post-mortem (min.)	specific-activity ratios × 1000		
				brain acid-sol. P to blood acid-sol. P	nucleoprotein P to brain acid-sol. P	phospholipid P to brain acid-sol. P
(a) age						
26	3·7	8		—	—	64·0
28	6·0	22		106	—	59·0
31	21·0	47		99	148	38·8
35	37·5	58	1·5	103	186	38·8
32	45·5			97	—	35·5
39	38·2	>365		132	190	42·3
40	40·1			118	160	41·2
(b) post-mortem autolysis						
46	26·7		4	100	146	41·6
14	34·3	60	6	—	112	39·8
22	33·3		10	115	98	36·4

Conditions as in table 3, except that for smaller mice, no. 26 and no. 28, the dose of radioactive phosphorus was 5 μcuries.

occurring before fixation of the brain tissue or by variations due to the age or sex of the animals.

The post-mortem changes occurring during standing of the brain after decapitation for 4 min. before fixation in trichloroacetic acid produced no significant change in the specific activity of the phosphorus in any of the fractions examined, but standing for 10 min. produced an apparent fall in the specific activity in the nucleoprotein fraction (table 2). In view of this observation, the time taken for the removal of the brain was reduced to a minimum and a standard time of approximately 1½ min. was allowed before fixation of the tissue.

It was to be expected that the turnover of phosphorus in the brain phospholipids would be greater in young animals than in old, and this has been shown to hold in the rat (Fries, Changus & Chaikoff 1940). In the present investigation, a series of mice of over a year old were compared with a litter of four young mice which were studied at various times up to 58 days after birth. The figures for the specific-activity ratio confirmed a higher turnover rate in the phospholipid fractions for the very young animals of 8 and 22 days; but the ratios were in the normal range for animals over 2 months in age (table 2). No significant differences between males and females in the specific-activity ratios were observed in these or in later experiments.

Normal series

The nine animals included in the normal series gave mean specific-activity ratios of 0·109, 0·156 and 0·0428 for the brain acid-soluble, nucleoprotein and phospholipid fractions respectively (table 3). In view of the fact that each individual ratio

TABLE 3. UPTAKE OF RADIOACTIVE PHOSPHORUS BY THE NORMAL MOUSE BRAIN

| | | | specific-activity ratios × 1000 | | |
mouse no.	weight (g.)	sex	brain acid-sol. P to blood acid-sol. P	nucleoprotein P to brain acid-sol. P	phospholipid P to brain acid-sol. P
1	37·6	m.	—	127	45·8
2	27·3	m.	—	151	46·4
3	33·3	m.	—	194	34·8
7	25·0	m.	140	127	45·9
10	28·0	m.	90	156	41·6
16	37·3	f.	100	138	43·8
19	—	f.	116	162	41·6
27	30·0	f.	91	183	43·5
42	30·0	m.	116	170	42·0
mean			109	156	42·8

The specific-activity ratios of the fractions were measured 3 hr. after intraperitoneal injection of 10 μcuries of radioactive phosphorus in 0·02 mg. sodium dihydrogen phosphate. The brains were ground with 10 % trichloroacetic acid within 1½ min. of decapitation.

depended on four estimations, two of which were done with the Geiger counter, the individual variation was relatively small. The variation was greatest for the main acid-soluble fraction and least for the phospholipid fraction, in which the results

for eight of the nine animals came within 10 % of the mean. The series therefore provided a satisfactory basis for further experiments in which the effects of various changes in the physiological state were studied.

THE EFFECTS OF ANAESTHESIA, REDUCED TEMPERATURE AND INSULIN HYPOGLYCAEMIA

Anaesthesia

A series of mice lightly anaesthetized with sodium pentobarbital ('Nembutal') during the 3 hr. experimental period gave lower activity ratios for the nucleoprotein and phospholipid fractions of the brain indicating a reduced phosphorus uptake in both of these fractions (table 4). The differences from the normal were statistically significant when tested by Fisher's t-test at the 5 % level of probability for the nucleoprotein and 2·5 % level for the phospholipid fraction.

TABLE 4. THE EFFECT OF (a) SODIUM PENTOBARBITAL (NEMBUTAL) ANAESTHESIA, AND (b) REDUCED TEMPERATURE ON THE UPTAKE OF RADIOACTIVE PHOSPHORUS BY THE MOUSE BRAIN

			specific-activity ratio × 1000		
mouse no.	degree of anaesthesia	body temp. (° C)	brain acid-sol. P to blood acid-sol. P	nucleoprotein P to brain acid-sol. P	phospholipid P to brain acid-sol. P
			(a) anaesthesia		
8			91	125	44·9
9	light	38	100	123	28·9
11			89	142	36·0
12			92	143	36·2
mean			93	133	36·5
		(b) anaesthesia and reduced temperature			
4			—	142	20·0
5	light	27	—	109	19·9
6			—	124	22·1
13	deep		54	55	1·8
mean			—	107	15·9

Anaesthesia was maintained for 3 hr. by intraperitoneal injection of sodium pentobarbital. At the start 50 mg./kg. was given, with further smaller doses if there were any signs of waking. The dose was increased for mouse no. 13 to give deep anaesthesia. Precautions were taken to maintain the body temperature in nos. 8, 9, 11 and 12; but it was allowed to fall in nos. 4, 5, 6 and 13. The conditions otherwise were as in table 3.

Reduced temperature

In the observations on the effects of anaesthesia the body temperature of the mice was carefully maintained during the experimental period by keeping them on a warmed operating table, so that the temperature was kept at 38° C. In a further series of experiments on anaesthetized animals the body temperature was allowed to fall to about 27° C. The combined effects of anaesthesia and reduced temperature

produced a considerably greater fall in the phosphorus exchange, as indicated by the specific-activity ratios for all three fractions. In one animal which was deeply anaesthetized the brain phospholipid activity ratio was 0·0018, indicating a fall to only about 4 % of the normal rate of phospholipid synthesis.

Insulin hypoglycaemia

The specific-activity ratio of the phospholipid fraction was greatly reduced by insulin hypoglycaemia: the mean ratio for five animals after injection of 0·15 unit of insulin was 0·0216, indicating a fall of approximately 50 % in the rate of synthesis of radioactive phospholipid (table 5). The insulin was injected just before the radio-

TABLE 5. (a) THE EFFECT OF ELECTRICALLY INDUCED CONVULSIONS, AND (b) THE EFFECT OF INSULIN ON THE UPTAKE OF RADIOACTIVE PHOSPHORUS BY THE MOUSE BRAIN

		specific-activity ratios × 1000		
mouse no.	remarks	brain acid-sol. P to blood acid-sol. P	nucleoprotein P to brain acid-sol. P	phospholipid P to brain acid-sol. P
		(a) electrically induced convulsions		
15	2 shocks of 15 sec. each (10 V, a.c.) at ½ hr. and 1½ hr. points	—	152	39·6
17	1 shock of 5 sec. (25 V, a.c.) at 10 min. point	106	152	37·1
18	1 shock of 9 sec. (15 V, a.c.) at 10 min. point	98	153	38·6
21	1 shock of 3 sec. (20 V, a.c.) at 3 min. point	—	140	36·9
48	1 shock of 1 sec. (25 V, a.c.) at 3 min. point	90	168	38·9
mean		—	153	38·2
		(b) insulin hypoglycaemia		
20	approx. 1 hr. coma	134	163	37·0
23	convulsions; coma approx. 2½ hr.	128	128	9·7
25	stuporose; no coma	108	—	31·4
29	approx. 1 hr. coma	114	134	19·0
47	approx. 1½ hr. coma	124	140	11·0
mean		122	141	21·6

Insulin (0·15 unit) was injected intraperitoneally just before the radioactive phosphate. The conditions otherwise were as in table 3.

active phosphorus and the mice were not allowed access to food. There was apparently a slight rise in the specific-activity ratio of the brain acid-soluble to blood acid-soluble phosphorus: this agrees with the observations of Kaplan & Greenberg (1944), who found that insulin hypoglycaemia caused a rise in the specific activity of the acid-soluble phosphorus fraction in the muscle and liver of the rabbit, while there was a fall in the specific activity of this fraction in the blood.

THE EFFECTS OF ELECTRICAL STIMULATION AND OF EXPOSURE IN A ROTATING DRUM

Electrical stimulation of the brain

Stimulation was carried out with stainless steel electrodes of 0·25 sq.cm. area, which were applied to the scalp 0·5 cm. posterior to the eyes. Contact was obtained by previously cutting the fur in this region and applying electrode jelly. The current used was 50 c./sec. a.c. at 10 to 25 V, and under the conditions used stimulation for 1 to 15 sec. generally produced convulsions after a latent period of about 10 sec. The spasms were severe for the first 10 sec. and then continued intermittently for about a minute.

The electrical stimulation produced no apparent change in the specific-activity ratios of the brain acid-soluble or nucleoprotein fractions; but there was a decrease of 11% in the ratio of the phospholipid fraction. The change was statistically significant at the 2% level of probability (table 5).

Rotating drum

The object of this series of experiments was to test if any measurable effect was produced on the phosphorus turnover in the brain by emotional excitement under normal physiological conditions. With this object, the mice were put for the experimental period of 3 hr. in a hollow perforated metal drum which was slowly rotated at $1\frac{3}{4}$ r.p.m. In this unaccustomed situation the mice showed the appearance of fear, as indicated by micturition, defaecation and occasional attempts to escape. They also took a certain amount of involuntary exercise, though at the slow rate of

TABLE 6. THE EFFECT OF EXPOSURE IN A ROTATING DRUM ON THE UPTAKE OF RADIOACTIVE PHOSPHORUS BY THE MOUSE BRAIN IN (a) NORMAL UNCONDITIONED ANIMALS, AND (b) IN CONDITIONED ANIMALS

| | | specific-activity ratios × 1000 | | |
mouse no.	duration (hr.)	brain acid-sol. P to blood acid-sol. P	nucleoprotein P to brain acid-sol. P	phospholipid P to brain acid-sol. P
		(a) normal unconditioned animals		
24	3	115	98	31·8
30	$2\frac{1}{2}$	100	146	27·1
33	approx. 2	108	179	25·6
34	approx. 2	132	—	29·0
43	3	123	188	39·5
45	3	110	164	37·6
mean		115	155	31·8
		(b) conditioned animals		
36	3	110	178	44·7
37	3	90	178	38·8
38	3	118	181	48·5
41	3	135	149	46·5
mean		113	171	44·6

The drum (diameter 18 in.) was rotated on a horizontal axis at $1\frac{3}{4}$ r.p.m.

rotation used, this was not very much greater than their normal spontaneous activity in their cages. To serve as a control, a series of littermates were conditioned to the rotating drum by being placed in it once or twice daily during the period of 3 weeks prior to the experiment; after a short time these animals showed no evidence of fear in the drum. The resulting specific-activity ratios showed no change in the acid-soluble or nucleoprotein fractions, but there was a fall of 25 % in the phospholipid activity ratio of the unconditioned mice exposed in the rotating drum. The change was significant at the 0·1 % level of probability. There was no significant change in the animals which had been conditioned to the rotating drum (table 6).

DISCUSSION OF RESULTS

The metabolism of the brain *in vivo* has been studied by determining the rate at which radioactive phosphorus is taken up into the different phosphorus-containing fractions of the mouse brain. Specific-activity ratios $\left(\dfrac{{}^{32}\text{P}/{}^{31}\text{P} \text{ of fraction } X}{{}^{32}\text{P}/{}^{31}\text{P} \text{ of fraction } Y} \right)$ were used as a measure of the rate of synthesis of radioactive phosphorus into the different fractions. Under carefully standardized conditions the individual variation in the ratios in the different animals was small, so that relatively slight changes in the rate of phosphorus uptake could be detected. This gave a sensitive method of observing the metabolic changes associated with variations in the state of functional activity of the brain *in vivo*.

The specific-activity ratio of the nucleoprotein fraction decreased during post-mortem autolysis of the mouse brain. If the fraction were chemically homogeneous, the decomposition of the nucleoprotein fraction would cause a rise in the ratio, owing to the liberation of phosphorus of relatively low specific activity into the acid-soluble fraction. The observed decrease in the ratio on autolysis suggests that the fraction contains one compound with relatively high turnover rate, which is decomposed by enzymic activity more rapidly than the rest of the fraction. It has recently been shown that the small quantity of phosphoprotein in this fraction has a much higher rate of turnover of radioactive phosphorus than the nucleic acids in a number of tissues (Davidson, Gardner, Hutchison, Macindoe, Raymond & Shaw 1949). Alternatively, the observed decrease in the ratio could be accounted for if the most recently laid down nucleoprotein molecules are the first to be broken down by enzymic degradation.

The specific-activity ratio indicating the uptake of radioactive phosphorus into the phospholipid fraction of the brain is largely independent of the specific activity of the acid-soluble phosphorus of the brain; but since the 'active' phospholipids formed during the 3 hr. experimental period are mixed with 'non-active' phospholipids, the specific activity of the total phospholipid fraction depends on the total phospholipid content of the brain. The specific-activity ratio of the phospholipid fraction was found to be higher in young animals than in old. This might be due to a higher rate of phospholipid synthesis in the laying down of new tissue. The total phospholipid concentration is, however, lower in the brain of the young animal, and this factor would also increase the ratio indicating radioactive phospholipid synthesis (Lang 1937).

In anaesthetized animals the rate of synthesis of radioactive phosphorus into the nucleoprotein and phospholipid fractions of the brain decreased with the general fall in the metabolic rate; this decrease was more marked when the metabolism was further lowered by allowing the body temperature to fall. In insulin hypoglycaemia and in electrically induced convulsions there was a significant fall in the phosphorus uptake into the phospholipid fraction without a corresponding change in the nucleoprotein fraction.

It was conceivable that the changes found in the specific-activity ratios could be due to an alteration in the circulation of the brain or in the permeability of the blood-brain barrier; but the results which showed a change in the brain phospholipids without affecting the nucleoproteins make it unlikely that the observed changes are due to an unspecific factor of that kind. The decrease in the uptake of radioactive phosphorus into the brain phospholipids in electrical stimulation and in emotional excitement is apparently a specific effect which may be attributed to a metabolic change affecting the rate of phospholipid synthesis.

A significant decrease of 25 % in the uptake of radioactive phosphorus into the phospholipid fraction of the brain was observed under normal physiological conditions in mice which were left for 3 hr. in a slowly rotating drum. This effect was entirely absent in a series of animals which had been accustomed to exposure in the rotating drum and which therefore showed no emotional reaction, such as excitement or fear, when put into it. The amount of muscular exercise taken during the experiment was not very great and was similar for the two groups: it therefore appeared that the observed changes should be attributed to the emotional excitement rather than to the exercise. This would agree with the observations of Le Page (1946), who was able to produce severe shock in rats by putting them in a rapidly rotating drum; but he found this effect was absent in rats which had previously been conditioned to the experience. It has been shown by an independent method that emotional excitement causes significant changes in other metabolites such as the lactic acid and acetylcholine in the brain (Richter & Dawson 1948; Richter & Crossland 1949).

The precise function of the phospholipids in nervous tissue is unknown, but it is known that they are conserved even in prolonged starvation and they have therefore been regarded as permanent or 'structural' elements of the cell. The relatively high turnover rate of the brain phospholipid fraction found in the present investigation gives evidence that in the brain, as in other tissues, the phospholipids are by no means metabolically inactive. The observed turnover rate was such that an amount of phosphorus equivalent to the total amount contained in the phospholipid fraction would be exchanged in 70 hr.

It has long been postulated that the physiological processes of facilitation and of memory may involve a change in the structural elements of the nerve cell. The question of the permanence of the observed changes in the phospholipid metabolism and their localization in the brain are problems for further investigation; but in the meantime these changes are suggestive in connexion with the problem of the 'engram', or the permanent traces left when nervous tissue is stimulated.

REFERENCES

Ashby, W. & Chan, D. V. 1943 *J. Biol. Chem.* **151**, 515.

Cohn, W. E. & Greenberg, D. M. 1938 *J. Biol. Chem.* **123**, 185.

Dameshek, W., Myerson, A. & Loman, J. 1934 *Amer. J. Psychiat.* **91**, 113.

Davidson, J. N., Gardner, M., Hutchison, W. C., Macindoe, W. M., Raymond, W. H. A. & Shaw, J. F. 1949 *Biochem. J.* **44**, xx.

Dawson, R. M. C. & Richter, D. 1950 *Amer. J. Physiol.* **160**, Jan. 1950.

Fiske, C. H. & Subbarow, Y. 1929 *J. Biol. Chem.* **81**, 629.

Fries, B. A., Changus, G. W. & Chaikoff, I. L. 1940 *J. Biol. Chem.* **132**, 23.

Gurdjian, E. S., Webster, J. E. & Stone, W. E. 1947 *Epilepsy*, p. 184. Baltimore: Williams and Wilkins Co.

Hevesy, G. 1939 *J. Chem. Soc.* p. 1213.

Hevesy, G. & Hahn, L. 1940 *K. danske vidensk. Selsk. Biol. Medd.* **15**, 5.

Jones, H. J. 1948 Personal communication in *Radioactive indicators*, p. 510. By Hevesy, G. New York: Interscience Publishers Inc.

Kaplan, N. O. & Greenberg, D. M. 1944 *Amer. J. Physiol.* **140**, 598.

King, E. J. 1946 *Micro-analysis in medical biochemistry*, p. 53. London: J. and A. Churchill Ltd.

Lang, A. 1937 *Hoppe-Seyl. Z.* **246**, 219.

Le Page, G. A. 1946 *Amer. J. Physiol.* **146**, 267.

Lindberg, N. O. H. & Ernster, L. 1950 *Biochem. J.* **46**, 43.

Maclean, H. 1914 *Biochem. J.* **8**, 453.

Richter, D. & Crossland, J. 1949 *Amer. J. Physiol.* **159**, 247.

Richter, D. & Dawson, R. M. C. 1948 *Amer. J. Physiol.* **154**, 73.

Schmidt, G. & Thannhauser, S. J. 1945 *J. Biol. Chem.* **161**, 83.

Tuttle, L. W., Erf, L. A. & Lawrence, J. H. 1941 *J. Clin. Invest.* **20**, 577.

Veall, N. 1948 *Brit. J. Radiol.* **21**, 347.

Does heat production precede mechanical response in muscular contraction?

By A. V. Hill, F.R.S., *Biophysics Research Unit, University College, London*

A decision between two current theories of muscular contraction is provided by the proof that the heat production in a twitch begins before the mechanical response. In the skeletal muscle of the tortoise at 0° C the heat production starts off at its maximum rate about 60 msec. after stimulation, while the tension cannot be detected before 90 to 100 msec. The heat of activation clearly precedes the mechanical response.

The processes of contraction and relaxation in a tortoise's muscle occur at about one-tenth of the speed found in a frog's muscle at the same temperature. This should allow much greater accuracy in various investigations which are limited at present by the time taken in manipulation or recording.

Two theories of muscular contraction are fundamentally opposed. In the first the mechanical response is supposed to result from chemical changes initiated by stimulation. In the second, mechanical energy is imagined to be released by the stimulus from some reservoir or accumulator previously charged. On the latter view, the liberation of heat, accompanying chemical change, would occur during the process of recharging; though some heat also might appear in the course of mobilizing the energy from the reservoir. Thus, on the second theory, heat could be liberated during and after the mechanical response, but not before it. On the first theory, however, heat might appear, as the consequence of chemical change, before any sign could be detected of the mechanical response. This heat would be associated with the process of activation (Hill 1949a).

A verdict between the rival theories could be reached if evidence were found that the heat production clearly preceded the mechanical response. In a recent paper (Hill 1949c) such evidence was presented. It was not, however, as conclusive as could be wished for so important a decision. The speed of recording the heat is limited by the time taken in heat-flow and in galvanometer response, and the true course of the early heat production could be determined only by a rather complicated analysis of the records. No doubt is suggested of the general validity of that analysis, but the great importance of the conclusion made it desirable if possible to confirm it more directly. No way being known of improving the speed of the equipment, the only thing to do was to use a slower muscle. At a given temperature the skeletal muscles of the tortoise (*Testudo mauretanica*) take about 15 times as long to complete the several phases of contraction and relaxation as do those of the frog (*Rana temporaria*); and, fortunately, they behave extremely well for long periods under experimental conditions.

The muscle finally adopted for the present purpose has been identified for me by Dr E. T. B. Francis of Sheffield University as the extensor iliotibialis. It was named the rectus femoris by Bojanus (1819–21, no. 99) and Owen (1866, fig. 151, no. 99). It is not unlike a frog's gastrocnemius, but it is very flat on one side, is

softer and more extensible and contains less tendinous material. Its flatness allows it to lie in very close contact with a thermopile, so avoiding harmful delay in heat-flow. It was stimulated by a condenser discharge of appropriate time constant applied near the entry of its nerve. A single muscle was used.

Hartree (1926) stated that he found it impossible to make satisfactory experiments on a tortoise muscle at 0° C. (He used the biceps cruris (Bojanus, Owen, no. 103) which Dr Francis tells me is now more properly described as the iliofibularis). The only difficulty found in the present investigation was that the heat production in a twitch was rather small (about 2 mcal./g.) and occurred so slowly that heat loss substantially diminished the later deflexion. Possibly Hartree used an unsuitable stimulus or an unsuitable frequency of stimulation (see below). The objective being to record the earliest appearance of heat, the instruments (see Hill 1949 b) were used with the highest available sensitivity and the results of several records were averaged together, each read to 0·1 mm., i.e. to about 10^{-6} °C.

The general character of the twitch is shown in the curves of figure 1, recorded at lower sensitivity and speed, allowance being made for heat loss which is important at the longer times. It is obvious even from figure 1 that the heat pro-
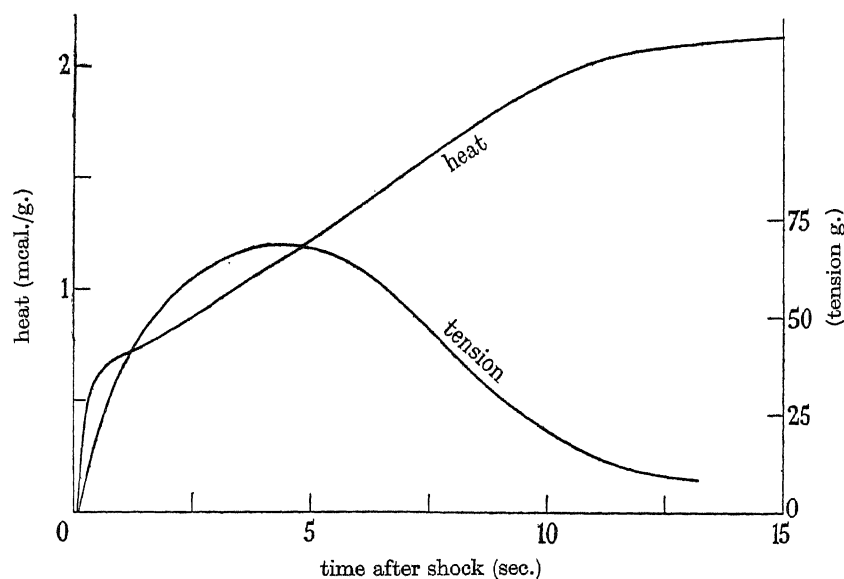
Figure 1. Tension and heat production in isometric twitch of tortoise muscle (175 mg., 25 mm.) at 0° C. Time after shock in seconds. Tension in grams weight. Heat in millicalories per gram of muscle. The heat has been corrected for heat loss.

duction begins appreciably before the mechanical response. In figure 2 the first 0·25 sec. of a twitch is shown on a much more extended scale. The tension was recorded photoelectrically, by the shadow cast by a stiff tension lever on a twin photocell, with a sensitivity which would have given a deflexion of 460 mm. for its maximum; it was read to 0·1 mm., i.e. to about 13 mg., every 10 msec. The heat was recorded with a sensitivity which would have given a deflexion of 213 mm. for the maximum, and the mean of five records is shown without any allowance or

analysis. In order to facilitate the comparison of heat and tension each is plotted as a percentage of its recorded maximum. The heat clearly precedes the mechanical response.
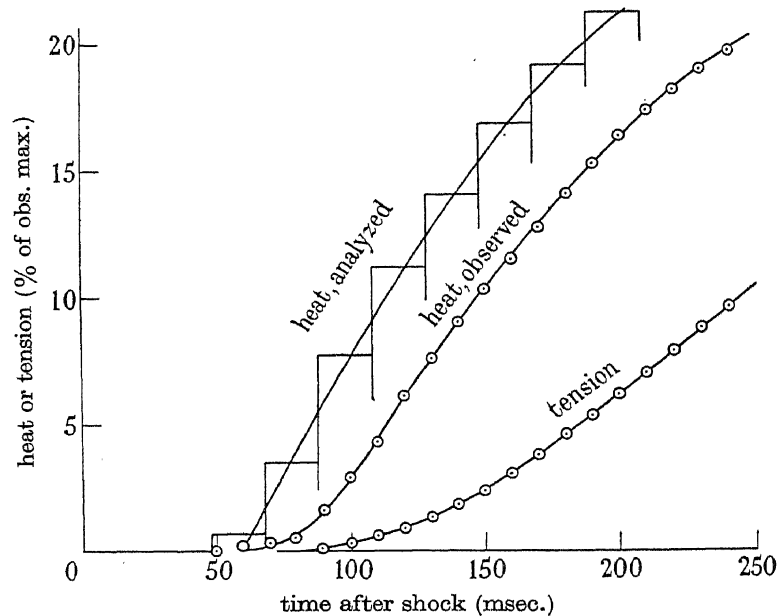
FIGURE 2. Tension and heat production in first 0·25 sec. of isometric twitch of tortoise muscle (160 mg., 25 mm.) at 0° C. Each is expressed as a percentage of its observed maximum. The observed maximum of the heat was about 25 % less than the true maximum, owing to heat loss; if this were allowed for, the scale for heat would be three-quarters of the numbers given. The heat curve is drawn from the average of five records read every 10 msec. from the start. The mean deviation of the circles from the smooth curve is about 0·1 % of the maximum deflexion, i.e. about 2×10^{-6}° C.

To the heat-curve of figure 2 an analysis was applied, as shown by the blocks, to allow for lag in heat-flow and in galvanometer response. The result indicates that the heat starts off at its maximum rate at about 60 msec., while the tension leaves the base-line tangentially, its first visible deviation being at 90 msec. The muscle was under sufficient initial tension to ensure that no slack had to be taken up and there was no significant cause of delay, such as friction or inertia, in recording the mechanical response; so the analyzed curve of heat and the observed curve of tension should be closely comparable. The gap between them at the start is so large that no doubt can exist that the heat starts substantially the earlier. This early stage of the heat is presumably the heat of activation (Hill 1949a); the heat of shortening accompanies and is proportional to the shortening (internal or external) which occurs later.

Several other experiments led to a similar result:

| start of heat (msec.) | 54 | 65 | 60 | 50 | 50 |
| start of tension (msec.) | 90 | 100 | 110 | 100 | 90 |

It appears, therefore, that at 0° C the heat production of a tortoise muscle starts about 60 msec. after a shock, while the mechanical response cannot be detected

before about 100 msec. In all these experiments the tension reached its maximum in about 4 to 5 sec. The corresponding time in a frog is about 0·3 sec., in a toad about 0·6 sec.; the frog, therefore, is 13 to 17 times quicker than the tortoise, the toad 7 to 8 times quicker. If all the times involved in contraction were in the same ratio, the beginning of heat production in the frog would be at about 4 msec., in the toad at about 8 msec. These are shorter than the times given in the earlier paper (Hill 1949c), namely, about 10 and 25 msec. respectively. Those times were deliberately not underestimated, and it may be that the shorter times derived by analogy from the tortoise are more nearly correct; though that is of little importance now, for the tortoise muscle has given a clear decision.

It may be, of course, that the piezo-electric method applied to record the tension in a tortoise muscle at 0° C would show a mechanical response starting appreciably earlier than 90 to 100 msec. In a frog's muscle at 0° C Abbott & Ritchie (1948) detected the initial negative phase at about 7 msec., the positive phase beginning at about 18 msec. In a muscle of one-fifteenth the speed, 7 msec. would become 105 msec., much the same as found in the present investigation. In comparing the thermal with the mechanical response, however, it is desirable to use comparable sensitivities. The sensitivity of the piezo-electric method is extremely high; if it detects the beginning of response rather earlier we must reflect that if the myothermic sensitivity could be increased in the same proportion it would probably detect the heat earlier too. With comparable sensitivities, as used in such experiments as that of figure 2, a truer impression is given of the comparative speeds of onset of the two aspects of response.

In the experiments (Hill 1949c) on the muscles of frog and toad, it was necessary to make allowance for the time taken in the propagation of the response along the muscle lying on the thermopile. Apparently, no such allowance was necessary for the tortoise muscle, possibly because the stimulating cathode lay close to the entry of the nerve, so that the whole muscle may have responded practically together. It was noticeable that a better and more rapid response was obtained in this way than with the cathode at the other end.

We may conclude from these experiments that the heat of activation precedes the mechanical response. This makes it impossible to accept, at least in its simple form, the theory that the mechanical energy for contraction is derived from some reservoir or accumulator where it remains in a latent form until released by a stimulus.

During this investigation the opportunity was taken of studying the time course of tension and heat production during a tetanus at 0° C. Figure 3 shows a record made on the same muscle as was used for figure 1. The tension took so long (more than half a minute) to reach its maximum (180 g.) that the heat loss was large; the allowance for it in 30 sec. at 3·5 %/sec., was 65 % of the observed deflexion. The stimulus was 1 condenser discharge every 1 sec. The tension developed was considerable, namely about 2·5 kg./sq.cm. of the average cross-section. The total heat increased approximately linearly with the duration of the stimulus, its rate being about 0·25 mcal./g. × sec. The opposite muscle of the same animal on the following day gave the same rate up to 50 sec., when the stimulus was stopped.

This rate is about one-tenth of the steady heat production during the maintained contraction of a frog's sartorius at the same temperature, about one-fifth of that of a toad's. In order to obtain a good contraction in a very slow muscle it is essential that the frequency of stimulation should be low; otherwise failure of the excitatory process occurs. It may be that Hartree's inability (Hartree 1926) to obtain results at 0° C was due to using too high a frequency of stimulation.
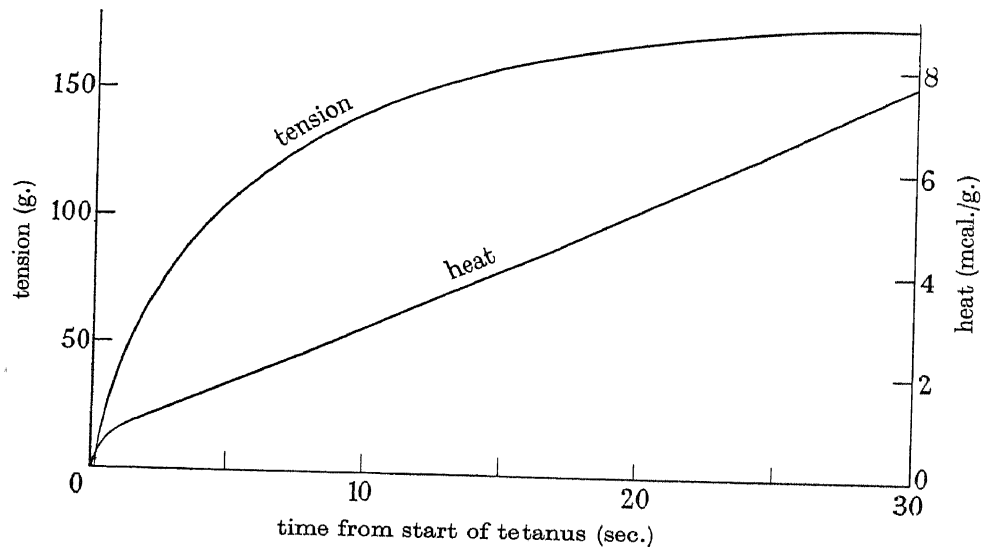
FIGURE 3. Tension and heat production during maintained contraction of a tortoise's muscle at 0° C (the same muscle as used for figure 1). Stimulus, 1 condenser discharge per sec. The heat deflexion has been corrected for heat loss (3·5 % per sec.).

The fact that the processes of contraction in the skeletal muscles of a tortoise at 0° C occur so slowly, taking 10 to 15 times as long as in a frog's muscle at 0° C or 60 to 90 times as long as in a frog's muscle at 20° C, provides a valuable means of observation by techniques which cannot be made fast enough for a more rapid muscle. For example, the measurement of pH changes inside a muscle fibre during contraction is subject to considerable delay due to the slowness of diffusion to the recording electrode (Hill 1948). The importance of this delay is reduced in inverse proportion to the time taken in the processes of contraction. Again, the accurate location in time of chemical processes supposed to take place during contraction and relaxation (Hill 1949e) would be made easier by using a slower muscle. The investigation of the abruptness of onset of the active state after a shock by the method of quick stretches (Hill 1949d) would be much more precise if applied to a muscle in which everything occurred at one-tenth of the speed. In a frog's muscle the liberation of energy is so rapid during a tetanus at room temperature that the contraction cannot be fully maintained for more than a few seconds. In a tortoise's muscle at 0° C, stimulated at a low enough frequency, a contraction can be kept up without sign of failure for many minutes.

Frog's muscles are generally used for physiological investigations because they are convenient and readily available; but they have the serious disadvantage, for many purposes, of very high speed. The tortoise's muscle is far slower, and behaves

extremely well for long periods under experimental conditions. Curiously enough (Hartree & Liljestrand 1926), the recovery heat production, for a given initial heat production, occurs at nearly the same rate in a tortoise's muscle as in a frog's. Although the processes of contraction and relaxation are so much slower, in accordance with the needs and habits of the animal, those of recovery are about the same. This may be one of the reasons why the isolated muscles of the tortoise behave so well for long periods. It would be interesting to find out whether the chemical processes which take place after contraction, in the presence or absence of oxygen, occur no more slowly than in frog's muscle, in spite of the great difference in speed of contraction and relaxation.

REFERENCES

Abbott, B. C. & Ritchie, J. M. 1948 *J. Physiol.* **107**, 5*P*.
Bojanus, Z. H. 1819–21 *Anatome Testudinis*. Vilna.
Hartree, W. 1926 *J. Physiol.* **61**, 255–260.
Hartree, W. & Liljestrand, G. 1926 *J. Physiol.* **62**, 93–97.
Hill, A. V. 1948 *Proc. Roy. Soc.* B, **135**, 446–451.
Hill, A. V. 1949*a* *Proc. Roy. Soc.* B, **136**, 195–211.
Hill, A. V. 1949*b* *Proc. Roy. Soc.* B, **136**, 228–241.
Hill, A. V. 1949*c* *Proc. Roy. Soc.* B, **136**, 242–254.
Hill, A. V. 1949*d* *Proc. Roy. Soc.* B, **136**, 399–420.
Hill, A. V. 1949*e* *Nature*, **163**, 320.
Owen, R. 1866 *Anatomy of Vertebrates*. Vol. I. London: Longmans Green & Co.

The series elastic component of muscle

By A. V. HILL, F.R.S., *Biophysics Research Unit, University College, London*

The properties have been examined of the undamped elastic component which lies in series with the contractile component of muscle. At higher tensions the elasticity is normal; the form of the load-extension curve as a whole must be largely due to the statistical distribution of tendon length in different fibres. The mechanical (elastic) energy of a contracting muscle is expressed graphically as a function of its tension. Even under completely isometric conditions this elastic energy is a significant fraction of the heat production in a twitch.

An important factor in the mechanical behaviour of muscle is the passive elastic component in series with the active contractile one (see Hill 1949*a*). This acts as a buffer when a muscle passes abruptly from the resting to the active state, and it accumulates mechanical energy as the tension of the muscle rises. If the muscle is opposed, as in most ordinary movements, by the inertia of a limb or an external mass, this mechanical energy can be used in producing a final velocity greater than that at which the contractile component itself can shorten. This is important in such movements as jumping or throwing. For simplicity in description the series elastic component will be referred to as 'tendon', but no assumption is implied that other undamped series elastic elements do not exist within the fibres themselves; the evidence of its properties is derived from mechanical experiments with active muscle, not from histological observation.

Some muscles have long tendons, or their fibre groups have long tendon bundles, while others have short ones. In the frog's gastrocnemius, for example (Hill 1931), the fibres run, on the average, only half the full length of the muscle; the other half must be tendon. The toad's sartorius, on the other hand (see Hill 1949 *b*), can shorten to about 30 % of its length in the body, which leaves little remainder for tendon. The former muscle is designed for rapid impulsive movements, the latter for slow crawling movements.

It is impossible to examine the properties of the series elastic component in a resting muscle; the contractile component at rest is so extensible that a load is taken almost entirely by the parallel elastic component, sarcolemma, etc. It is necessary to work with an actively contracting muscle, in a range of lengths within which the tension of the parallel elastic component can be neglected. Static experiments are impossible, because (*a*) a muscle fatigues too rapidly, and (*b*) because its active component shortens if the tension is lowered. It is necessary, therefore, to use a dynamic method, recording the tension as a function of length while shortening occurs at a speed greater than the maximum speed of the contractile component.

If a muscle is stretched while actively contracting a large part of the work done on it vanishes, that is to say, does not appear either as heat or as elastic mechanical energy; presumably it is used in driving some endothermic chemical process backwards. The muscle resists strongly and its tension rises to a high value during the stretch. In studying the time relations of this absorption of work it was necessary to know approximately how much of the work at any moment was still present as elastic energy in tendons and recording system—none was present after relaxation because the tension had fallen to zero. To avoid unnecessary complication it was advisable to make this elastic energy as small as possible, so that the tension recorder and its connexions to the muscle were made as inextensible as was consistent with a small enough inertia and a sufficient sensitivity, and a minimum of tendon and connecting thread was allowed at the end of the muscle. Nevertheless, substantial elastic energy was present in the system under the high tensions developed during the stretch of an actively contracting muscle. It was determined in the way described below. It should be noted that the results given here refer to two types of muscle only, the sartorii of the English frog and English toad, with a minimum of tendon at the tibial end. With other muscles the quantities found might be considerably larger; the methods described, however, are applicable to any muscle.

A Levin Wyman ergometer (Levin & Wyman 1927) constructed by Messrs C. F. Palmer was used. It had a magnetic release controlled by a key of a moving contact breaker, other keys of which determined the stimulus. To secure quicker attainment of the final velocity the driving weight was replaced by a powerful spring. Oil as the damping fluid was changed for silicone (D.C. fluid 200; 500 or 200 centistokes) which has a smaller temperature coefficient and is clean and reliable. The recording system consisted of a spring tension-lever carrying a vane of blackened aluminium foil which cast a shadow on a twin vacuum photocell. The output from this was transferred to an amplifier and recorded on one beam of a double beam cathode-ray tube; in the myothermic experiments the other beam was used for the heat. The tension lever, photocell and lamp were carried on the moving arm of the ergometer.

The start and end of the movement, after release, were recorded by contacts as slight flicks on the tension record. The tension lever was joined to the muscle by a light chain, which was adjusted to be exactly vertical in order to avoid sagging at low tensions.

A frog's sartorius was tetanized isometrically at 0° C with 5 maximal condenser discharges each way per second, and after 1·5 sec., when it had developed its full tension, it was released 5 mm. with the Levin Wyman ergometer at a speed of 42 to 46 mm./sec. The maximum velocity of shortening of the contractile component is about $\frac{4}{5}$ times the length of the muscle per second (Hill 1938); the muscles usually being 26·5 to 28·5 mm. long, their maximum velocity would be 35·5 to 37 mm./sec., so that their tension fell to zero after about 3·5 mm. release. The initial length was less than the resting length in the body, so that the initial resting tension was very small; after 3·5 mm. shortening it was negligible.

The equation of motion of the ergometer accelerating after release is

$$s = v_0 \left\{ t - \frac{Mv_0}{F} (1 - e^{-tF/Mv_0}) \right\},$$

where s is the distance travelled, v_0 the final velocity, t the time after release, M the mass, and F the driving force provided by the spring. M, F, s and v_0 are expressed in equivalent values at the point of attachment of the chain from the muscle to the tension recording lever. F was large compared with the force exerted by the muscle and M/F was found to be about 10^{-3} sec.^{2} cm.$^{-1}$. After acceleration was complete s became $v_0(t - Mv_0/F)$, so that a total distance lag of Mv_0^2/F finally occurred. With $v_0 = 4·5$ cm./sec., $Mv_0^2/F = 2 \times 10^{-2}$ cm. or 0·2 mm. This is small, and could be allowed for by extrapolating the falling tension curve backwards, as soon as it became linear (which it did, see below), to the initial level of tension. It would be better to use lower velocities, so that Mv_0^2/F would be less; but this would permit a greater total contractile shortening of the active component of the muscle during the release. In order to reduce the contractile shortening it would be better to make v_0 greater, but this would increase the lag during acceleration; and it had the further disadvantage of setting up more vibration in the instrument. If F could be increased the lag would be diminished, but it was already 2·2 kg., about as much as the ergometer and its release mechanism could safely stand. An ergometer could be designed which would have a considerably smaller M, but its construction would take time and effort better expended otherwise. The velocity of about 4·5 cm./sec. was chosen as the best compromise.

The recording system is extensible, consisting of thread, chain, tension lever, etc., and its contribution to the shortening observed in the muscle experiments was by no means negligible; indeed, if care was not taken it could make up a large part of it. It is easy, however, to allow for the extensibility of the recording system alone by a blank experiment without a muscle, the chain being joined to a fixed support in the usual muscle chamber by a thread of the same length as was used with a muscle. By adjusting the recording system with a screw, the tension was raised, and then the ergometer was released and a tension-shortening curve recorded exactly as in the muscle experiments. From this the shortening of the recording system and its connexions, between any two tensions, could be read off and subtracted from the shortening observed in the muscle experiments between those

tensions. In this way the true relation was determined between tension and shortening of the muscle itself, without any error due to the elasticity of the recording system and connexions.

The shortening of the muscle so determined is made up of two parts; (a) the shortening of the series elastic component, and (b) that of the contractile component. Of these, (a) is what we wish to determine, so that (b) must be allowed for. In principle, (b) could be eliminated by using a release so quick that the contractile shortening was negligible in the short time occupied by it. This would require a velocity of release much faster than that actually employed, and the equipment would not permit it. Even if special apparatus were constructed for the purpose, a further difficulty might arise from the natural viscosity of the series elastic elements. At the speeds actually used these elements are sufficiently nearly undamped; at much higher speeds even a small viscosity might diminish the force exerted externally during shortening. It was necessary, therefore, to allow for the contractile shortening.

According to the characteristic equation relating force and speed,

$$v = b(P_0 - P)/(P + a),$$

where P is the force exerted by the contractile component shortening at velocity v, P_0 is the isometric tension, and a and b are constants. Of these, P_0 was the starting tension in each experiment, and a was taken as $\frac{1}{4}P_0$, b as $\frac{1}{5}l_0$ per sec., the mean values found for the frog's sartorius at $0°$ C (Hill 1938). To apply the equation, the force-shortening curve was divided into equal intervals of time δt, and P was taken as the mean force in each interval. Then the amount of contractile shortening in any interval, $\delta s = v\delta t = b\delta t\,(P_0 - P)/(P + a)$, was calculated and so the total amount of shortening from the start was obtained by addition.

This calculation could not be exact, because it was necessary to assume mean values of a and b; it was not practicable to determine them directly in each experiment, for that is a fairly lengthy procedure and might alter the condition of the muscle for the subsequent observations. The error, however, cannot be large. The only way to avoid it completely would be to work with very quick releases, which might introduce (as suggested above) another and worse error.

The allowances having been calculated for (a) the extensibility of the chain and lever system and (b) the contractile shortening, these were added together and subtracted from the observed shortening. The result is shown in figure 1, for the mean of a number of records made on four muscle pairs freshly prepared from frogs in very good condition. Other experiments gave similar results. For generality, and so that the curve can be applied to any frog's sartorius at $0°$ C whatever its length and weight, the tension is expressed as Tl_0/M, which is very nearly the same as the force per sq.cm. of the muscle's cross-section (that is, $\rho Tl_0/M$, where ρ is the density of the muscle), and the shortening is expressed as a fraction of the resting length in the body.

At tensions above a Tl_0/M value of 600 a normal elasticity (Δ extension $\propto \Delta$ load) is shown in figure 1. The form of the curve below that must be due largely to statistical factors. In a large population of muscle fibres, each with its own tendon, it is very unlikely a priori that the lengths of these tendons would be the same. If x be
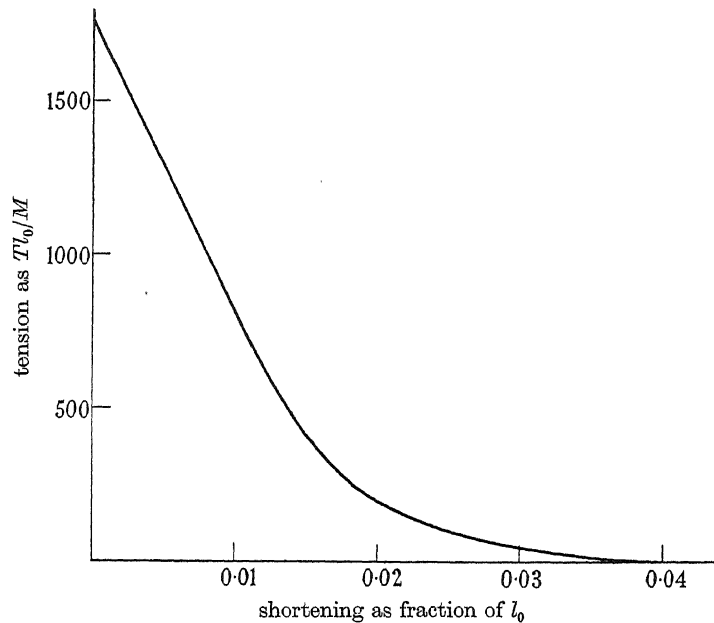
FIGURE 1. The relation between force and shortening during the rapid release of a frog's sartorius at 0° C during a maximal maintained contraction. Mean of a number of records obtained with four muscles. Allowance has been made (1) for the elastic shortening of the recording system and (2) for the contractile shortening of the active muscle. For generality, the tension is expressed as Tl_0/M, where M is the mass of the muscle, and the shortening as a fraction of the length l_0 at rest in the body (see text).

In order to average the results, the shortening was measured, plus or minus, from the length at which Tl_0/M was 1000 in each case. It is not possible to define the length of zero tension and the maximum tension varied from muscle to muscle. The four muscle pairs of figure 1 were:

(1) 28·5 mm. long, 185 mg.
(2) 28·5 mm. long, 202 mg.
(3) 27·5 mm. long, 202 mg.
(4) 26·5 mm. long, 149 mg.

The units of shortening chosen were $0\cdot0035\ l_0$ (about 0·1 mm.)

shortening units	0	1	2	3	4	5	6	7	8	9	10	11
Tl_0/M muscle (1)	1530	1270	1000	740	481	250	159	97	54	31	15	0
Tl_0/M muscle (2)	1610	1320	1000	660	378	217	129	74	31	13	0	0
Tl_0/M muscle (3)	1670	1310	1000	685	389	245	163	95	55	29	10	0
Tl_0/M muscle (4)	1840	1410	1000	598	364	242	155	98	66	44	27	12
Mean Tl_0/M	1660	1335	1000	671	403	238	151	91	21	29	13	3

It was very important that the muscles should be in good condition and uninjured. If one region were not contractile it would act as a series elastic body and the extensibility found would be much greater.

the stretched length and x_0 the resting length of a tendon, then the total force exerted is $\Sigma c\ (x - x_0)$, where c is an elastic constant assumed for simplicity to be the same for each. If the length of the muscle as a whole be diminished by y, the force becomes $\Sigma c\ (x - y - x_0)$, so long as all the fibres remain tight. If some of them become slack when the length has been diminished by y the force is $\Sigma c\ (x - y - x_0)$ for the remainder. Thus $dT/dy = cn/N$, where n/N is the fraction of fibres still taut

after shortening y. If we suppose n/N to follow the normal distribution shown by curve A of figure 2, then the value of T can be obtained by integration, being proportional to the area under the curve starting from the right. Curve B gives the result, and is obviously similar to the curve of figure 1. Conversely, by measuring the slope $(-dT/dy)$ at any point of the curve of figure 1 we can obtain the value of n/N at that point.
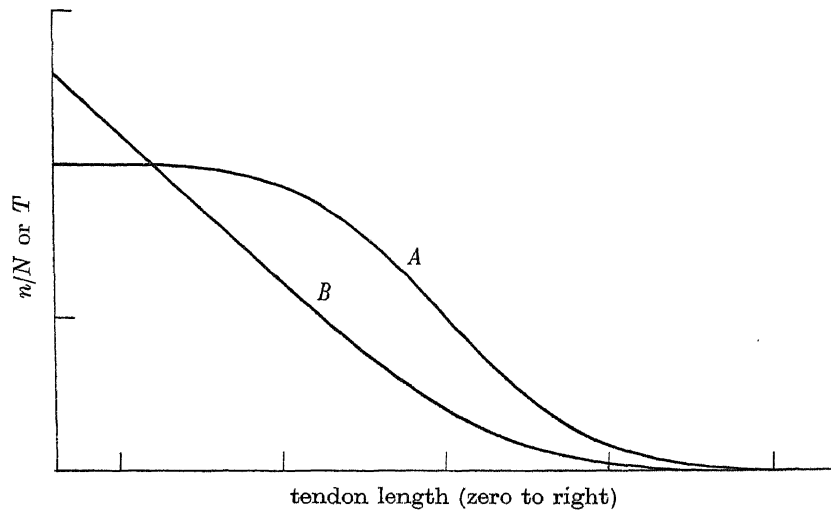
FIGURE 2. The statistical nature of the curve of figure 1. Curve A, n/N the fraction of the fibre tendons which are just tight at any length, assumed to follow a normal (Gaussian) distribution. Curve B, the total tension exerted by all the fibres that are tight at any length (see text).

ery likely there are other factors affecting the form of the relation in figure 1, but it is difficult to imagine an assembly of several hundred fibres which did not have a statistical distribution of lengths of their tendons similar to that shown in curve A, figure 2, and this must necessarily produce a tension-extension curve of the type found.

The tension-extension curve of the series elastic component has not been directly determined before. In a recent paper (Hill 1949a) it was calculated from the form of the isometric contraction of a toad's sartorius, but the result necessarily included the extension of the connexions of the muscle to the tension lever and of the lever itself, which was not separately determined. It must be admitted that in the past sufficient attention was not paid to the extensibility of the recording system and its connexions to the muscle; so that the extensibility of the series elastic component of the muscle itself was overestimated. Experiments with abrupt releases from maximal tension appeared to show (Gasser & Hill 1924) that about 10 % release was necessary to reduce the tension to zero, but control releases from the same tension, of the recording system alone without the muscle, were not made. If they had been, and if the muscles were in good condition contracting throughout their length, the amount of abrupt release required to reduce the tension of the muscle to zero would have been found considerably less.

The original purpose of the present experiments was to determine how much mechanical energy was present in a muscle and its mechanical recording system,

due to elastic stretch. The elastic energy in the muscle itself is obtained by integration from the curve of figure 1. It is shown as a function of tension in the curve (not the circles) of figure 3. For generality, the elastic energy W $(= \int T dx)$ is expressed as W/M, where M is the mass of the muscle, and the tension as Tl_0/M, where l_0 is the resting length in the body. The following example shows how figure 3 can be used. A sartorius 3 cm. long weighing 0·15 g. has developed a tension of 80 g., so $Tl_0/M = 1600$. From figure 3, $W/M = 15·6$, so that $W = 2·34$ g.-cm. If this were transformed into heat in the muscle (e.g. in relaxation) the rise of temperature would be $4·2 \times 10^{-4}$ ° C, which is about one-seventh of the rise of temperature in a single twitch.

When a muscle is stretched during contraction, the tension with which it resists the stretch may be considerably greater than that which it can exert isometrically. In estimating its elastic energy, therefore, it may be necessary to go beyond the
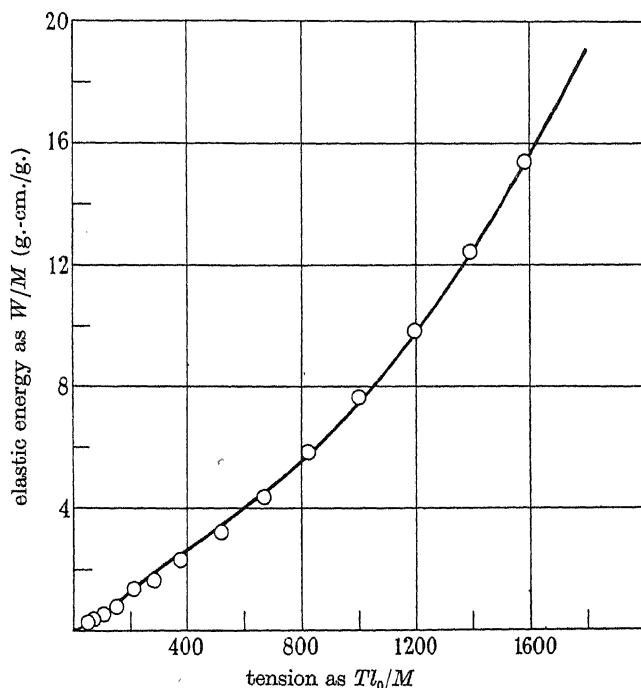
FIGURE 3. The elastic energy W of a muscle as a function of its tension. For generality the elastic energy is expressed as W/M (g.-cm. per g.) and the tension as Tl_0/M. The curve was calculated for frogs' sartorii from figure 1. The circles were calculated from similar measurements made on toads' sartorii.

range of figure 3. If we assume, as figure 1 strongly suggests, that the elasticity is normal above about $Tl_0/M = 600$, the curve from there onwards can be taken as $W/M = 2·0 + 5·3 \ (Tl_0/M)^2$.

In myothermic experiments it is necessary to have the muscle in a chamber in a thermostat rather far removed from the mechanical recording system outside. The connexion between them, 35 to 40 cm. long, is best made with a light chain. The chain must be light in order to avoid inertia, so is bound to be slightly extensible. With the best chain available (of the type formerly used with pince-nez) and

the present tension lever, the mechanical (elastic) energy present in the recording system was determined as a function of tension by the methods described above. The result was expressed in a graph similar to figure 3. The muscle referred to in the previous paragraph but one was supposed to have developed a force of 80 g. At this tension, from the graph, the mechanical energy in the recording system was 1·12 g.-cm., which is about one-half of that in the muscle itself. Together they make up 3·46 g.-cm., and if this were transformed into heat in the muscle the rise of temperature would be $6·2 \times 10^{-4}$ °C, about one-fifth of that in a muscle twitch.

The recording system was a good one; a more ordinary tension lever and connexions to the muscle, being more extensible, might very well contain several times as much elastic energy under a given tension. If it contained 4 g.-cm. in the above example the total mechanical energy would be 6·34 g.-cm., which (transformed into heat in the muscle) would lead to a rise of temperature of $1·14 \times 10^{-3}$ °C, nearly 40 % of the rise of temperature in an ordinary twitch. These quantities show the importance of the elastic energy developed by a muscle under supposedly 'isometric' conditions, and explain the relatively large amount of heat always found in relaxation after an 'isometric' contraction.

The measurements referred to hitherto were made on the sartorii of frogs (*Rana temporaria*). Three experiments, however, were performed in exactly the same manner and for the same purpose on the sartorii of toads (*Bufo bufo*), and the results are shown by the circles in figure 3. These muscles were:

(1) 24·5 mm. long, 102 mg.

(2) 23·2 mm. long, 64 mg.

(3) 25·5 mm. long, 74 mg.

The speed of release of the ergometer was 35 mm./sec., which in relation to the maximum speed of contractile shortening of toads' muscles is relatively higher than in the frog experiments. The corrections for (a) elastic stretch of the chain and recording system and (b) contractile shortening during release, were made exactly as for frogs' muscles, with the only difference that b was taken as $l_0/4·5$ per second instead of $l_0/3$, corresponding to the slower shortening of toads' muscles.

The agreement shown in figure 3 is good—indeed, the solid curve for frogs might have been drawn to fit the circles for toads. The toads' muscles were rather shorter and much smaller, and the agreement confirms the advantage of expressing the relation generally in terms of W/M and Tl_0/M. The experiments on frogs and toads were made independently and 5 months apart (the former in November, the latter in June), and the coincidence of the results would have been satisfactory even if they had been obtained on the same animal. The agreement between the different animals is striking.

REFERENCES

Gasser, H. S. & Hill, A. V. 1924 *Proc. Roy Soc.* B, **96**, 398–437.
Levin, A. & Wyman, J. 1927 *Proc. Roy. Soc.* B, **101**, 218–243.
Hill, A. V. 1931 *Proc. Roy. Soc.* B, **109**, 267–303.
Hill, A. V. 1938 *Proc. Roy. Soc.* B, **126**, 136–195.
Hill, A. V. 1949 a *Proc. Roy. Soc.* B, **136**, 399–420.
Hill, A. V. 1949 b *Proc. Roy. Soc.* B, **136**, 424–431.

A discussion on the action of local hormones

(*Discussion held* 16 *February* 1950—*Received* 17 *March* 1950)

INTRODUCTION

The original subject suggested for this discussion was the mechanism of the action of drugs. There are, of course, many ways in which drugs act, and so wide a subject might have prevented any close consideration of progress made in a more limited field. Certain drugs, such as ephedrine, eserine, atropine and others, are known to act in relation to adrenaline and acetylcholine which are active substances present in the body. The action of other drugs, such as quinine, procaine, pethidine and antihistamine agents, is also beginning to find an explanation in their relation to adrenaline, acetylcholine and histamine.

It therefore seemed suitable to arrange this discussion around our new knowledge of these substances and to call the discussion 'The action of local hormones'.

The suggestion in the title is not new. Feldberg & Schilf (1930) spoke about 'tissue' hormones, and Gaddum (1936), in his book, spoke of substances which play an important role by regulating the activity of tissues locally. Since that time attention has been mainly concentrated on adrenaline and acetylcholine as humoral transmitters and little has been said of their function as local hormones. Now, however, attention is turning to this function. We are curious to know why histamine should be present in such large amounts in nerve as v. Euler has shown, why acetylcholine should be present in placenta (Chang & Gaddum 1933) and why it should be present in and synthesized by trypanosomes (Bülbring, Lourie & Pardoe 1949). In one case, which I propose to describe in more detail, we have been able to obtain suggestive evidence concerning the part played by

acetylcholine in the auricles of the heart. This is a good example because it indicates clearly what is meant by the term local hormone.

The usual effect of acetylcholine on the heart is inhibition. Statements that acetylcholine could also stimulate the heart in certain conditions were made by Sachs (1937) and later by other workers (Spadolini & Domini 1940; Prosser 1942). Abdon (1944) described the preparation of an extract from rabbit heart which he said must contain acetylcholine precursor, because on treatment with HCl, acetylcholine was set free. He expressed the view that acetylcholine must play a role in the contraction of cardiac muscle.

The action of paludrine

The difficulty in accepting this view was obvious. The action of the vagus nerve, and of acetylcholine which is liberated when the vagus nerve acts, is not to promote cardiac contraction but to arrest it. Acetylcholine in the main acts as a brake on the heart, and it is not easy to understand how the brake can also drive the engine. Some light was shed on this, however, when my colleague Vane investigated the action of the antimalarial substance paludrine. In the course of this work he examined the effect on the spontaneously beating auricles, and found that under the influence of paludrine the contractions of the auricles were slowly diminished and the inhibitor action of acetylcholine was by stages converted into a stimulant action. Eventually the auricles ceased to beat, and the observation was made that the contractions began again when acetylcholine was added to the bath; the contractions were arrested once more when acetylcholine was removed. Thus by the action of paludrine on the auricles acetylcholine had been transformed from something which acted as a brake on the heart into something which would start its contractions (Burn & Vane 1949).

Acetylcholine on exhausted auricles

The evidence thus obtained was only in part satisfactory because it involved the use of paludrine, the action of which was obscure. My colleague Dr Bülbring was successful (Bülbring & Burn 1949a) in demonstrating that the same change in the action of acetylcholine on the auricles could be produced without using paludrine by allowing the auricles to beat in oxygenated Tyrode solution until the contractions finally disappeared at the end of 24 to 36 hr. The addition of acetylcholine again started the beat, and a further addition made the beat still more vigorous. When the beat was fully established, the addition of acetylcholine once more caused inhibition.

These observations made it very likely that acetylcholine played a part in auricular contractions and suggested that the contractions could occur only so long as the tissue was capable of synthesizing it.

Acetylcholine synthesis

The means of testing the suggestion lay at our disposal. Feldberg & Mann (1946) had investigated the synthesis of acetylcholine by a powder prepared from brain, and Comline (1946) had shown that a powder similarly prepared from auricles would also synthesize acetylcholine.

The method of Feldberg & Mann consists in preparing the powder by extracting the finely subdivided tissue in ice-cold acetone. The dried material is then incubated with choline and citrate in the presence of ATP. Fluoride and eserine are added to preserve the acetylcholine formed.

The following observations were then made:

(1) Fresh auricles synthesized about $45\,\mu g./g.$ powder/hr.

(2) Auricles which had ceased to beat after 24 to 36 hr. synthesized much less, about $10\,\mu g./g./hr.$

(3) Auricles which beat again on the addition of acetylcholine synthesized about $35\,\mu g./g./hr.$

These findings showed a relation between contraction and power to synthesize acetylcholine. The relation was, however, made much clearer by further observations on the effect of acetylcholine on the synthesis. It was found:

(4) That just as acetylcholine added to the bath containing the actively beating fresh auricles depressed the beat, so acetylcholine added to the acetone powder prepared from fresh auricles depressed the synthesis.

(5) That just as the addition of acetylcholine to the bath containing auricles which had ceased to beat caused the beat to start again, so acetylcholine added to the acetone powder prepared from stopped auricles augmented the synthesis.

Thus we had evidence of a direct relation between contraction and acetylcholine synthesis in the first place, and evidence that the changes in the activity of auricles in the bath produced by the addition of acetylcholine were due to change in its synthesis in the tissue. In this preparation we were able to interpret the changes seen on the smoked drum in terms of a biochemical process.

Acetylcholine as a local hormone

The position of acetylcholine as a local hormone now becomes clearer. We are presumably all agreed that the fundamental mechanism of the contraction in cardiac muscle depends on actomyosin activated by adenosine triphosphate. In skeletal muscle we know that the mechanism remains quiet until the liberation of acetylcholine by the motor nerve causes contraction. We now have a suggestion that in cardiac muscle the mechanism for firing off the contraction is also acetylcholine, but instead of being liberated by a nervous impulse, it is synthesized and causes a contraction probably when a certain concentration is reached. It is possible that the pacemaker controls the rate of beating by controlling the rate at which this concentration is reached. I realize that I am running into danger in attempting to be too precise in defining the function of the synthesized acetylcholine. However, the idea that the stimulus to contraction in the auricles is acetylcholine is tempting, and certainly it appears that the exercise of some such function is the task of acetylcholine as a local hormone. We at once understand why the action of quinine in skeletal muscle is so similar to that of quinidine in cardiac muscle. The function of acetylcholine is very similar to that of a humoral transmitter, but exercised in tissues where the activity is (at least in part) independent of central nervous control.

The action of drugs

The auricles provide an example to illustrate the action of drugs in relation to that of local hormones. As already described the application of paludrine produces arrest of the auricles and the same reversal of the action of acetylcholine as occurs in the exhausted auricles.

How then does paludrine act? The first possibility is that it reduces the ability of the auricles to synthesize acetylcholine. This is not so. A powder prepared from auricles which have ceased to beat under the influence of paludrine is capable of synthesizing acetylcholine as well as a powder from a fresh auricle. Thus the acetylcholine is synthesized but in the presence of paludrine cannot exert its usual effect.

Some light is shed on the problem by considering the effect of paludrine on that other form of striated muscle, the rectus abdominis of the frog. In the highest concentrations paludrine has a stimulant action of its own. In moderate concentrations it has no action of its own, but it diminishes the stimulant action of acetylcholine. To obtain the same stimulant action a much larger concentration of acetylcholine is required in the presence of paludrine. These effects suggest that paludrine has the power of combining with receptors with which acetylcholine combines when it acts, and that paludrine is a substance which competes with acetylcholine. This idea is easily applied to the auricle; we suppose that paludrine causes a steady decrease in the amplitude of contractions by attaching itself to more and more of the receptors so that the access of acetylcholine to these is increasingly blocked. The addition of acetylcholine no longer inhibits since it does not produce excess, but it augments the contractions by raising the total concentration of acetylcholine and so displacing some of the paludrine molecules. When the auricles finally stop, the acetylcholine synthesized is completely blocked by paludrine, but additional acetylcholine raises the concentration to a point at which the beat will start again in some of the fibres. Here then is an example of the mechanism of action of a drug. It acts by competition with a local hormone, in this case acetylcholine.

The blood vessels

We have at the present time no evidence for any other tissue similar to that obtained for the auricles, and ideas concerning the part played by local hormones in other tissues contain much more speculation. A little can be said about the blood vessels, however, which resemble the auricles in one respect, namely, that the normal action of acetylcholine on them in the body is inhibitor. If the inhibition is caused in the same way in the blood vessels as in the auricles, then it may be that a synthesis of acetylcholine can occur in the blood-vessel wall, and that the effect of additional acetylcholine in causing inhibition is again an effect of excess. One or two preliminary experiments have shown that an acetone-dried powder prepared from guinea-pig aorta can synthesize acetylcholine. Moreover, certain observations made on the vessels of the rabbit ear have a resemblance to the observations made in auricles (Bülbring & Burn 1949b). Thus when the ear vessels are perfused with Locke's solution, the initial reaction of the vessels to acetylcholine is a dilatation, that is to say, an inhibition. When the perfusion is continued overnight, the reaction

changes, and the same amount of acetylcholine then causes constriction or stimulation. Thus we can suppose that the initial inhibitory effect is the effect of excess of acetylcholine, while the later constrictor effect is seen when the normal rate of synthesis of acetylcholine in the vessel wall is much reduced by prolonged perfusion with Locke's solution. But the perfusion also removes diffusible substances from the blood-vessel wall, and it may be that these include other local hormones. Schmiterlöw (1948) has shown that the arteries of the ox contain not only acetylcholine, but also histamine and noradrenaline. Perhaps the change from the dilator to the constrictor effect of acetylcholine is due to the removal of histamine. I have found that the continuous perfusion of histamine in the Locke's solution in a concentration of $10\,\mu g./ml.$ will first increase the constrictor action of acetylcholine and then convert it to a dilator action once more.

ON THE ORIGIN AND FUNCTION OF THE ACETYLCHOLINE IN THE INTESTINAL WALL

By W. S. FELDBERG, F.R.S.

This discussion is intended to show that the substances released at nerve endings, the so-called chemical transmitters or mediators, may have additional physiological functions independent of nervous activity. During the last two years Dr Lin and I (Feldberg & Lin 1949, 1950) have carried out experiments on the spontaneous release of acetylcholine from the wall of the digestive tract, which appears to be an example of an acetylcholine metabolism of non-nervous origin.

Historically we have to go back to the old experiments of Weiland (1912), Le Heux (1918–19) and Magnus (1920), who discovered that large amounts of choline are continuously released from the intestinal wall. They, in fact, looked upon choline as the hormone of intestinal movements. Later it was shown that not only choline but acetylcholine as well is released from the intestinal wall (Feldberg & Rosenfeld 1933; Donomae & Feldberg 1934), and that this release continues after degeneration of the extrinsic nerves to the intestine (Bacq & Goffart 1939). At that time the nerve cells of the myenteric and submucous plexus were considered to be the source of the acetylcholine release. The results of our experiments, however, suggest that the nerve cells in the intestinal wall are probably not responsible for the acetylcholine release. These experiments are concerned with (1) the effect of cocaine on the release of acetylcholine and (2) the distribution in the intestinal wall of the enzyme responsible for its synthesis, the choline acetylase.

Cocaine and acetylcholine output

If we perfuse a sympathetic ganglion with saline solution, which must contain eserine to prevent the destruction of any acetylcholine which might be liberated, then acetylcholine appears in the venous effluent on stimulation of the preganglionic nerve fibres; but not if the saline solution contains in addition cocaine, about 1 in 10,000 or 20,000, which blocks nervous activity. But the spontaneous release of

acetylcholine from a piece of intestine perfused from its artery is not prevented by cocaine. Table 1 gives the amounts of acetylcholine released during 1 hr. perfusion from pieces of 100 cm. length of rabbit's small intestine, in the absence and presence of cocaine in the perfusion fluid.

TABLE 1. COCAINE ON ACETYLCHOLINE OUTPUT FROM 100 CM. PERFUSED RABBIT INTESTINE (FROM FELDBERG & LIN 1949)

cocaine concentration	acetylcholine output in μg./hr.						
0	1·5	1·8	2·4	3·5	4·5	5·2	5·5
1:50,000	5·8						
1:25,000	1·4	5·8					
1:1,000	2·0	6·7					
1:800	3·5	6·1					

A concentration of cocaine 1:800 to 1:1000 is much stronger than that required to inactivate the nervous activity in the intestinal wall. This can be shown when testing the effect of cocaine on the nicotine response of the isolated intestinal preparation. As seen from figure 1, the presence of 400μg. cocaine in the 15 ml. bath practically abolishes the strong stimulating action of 30μg. nicotine on the guinea-pig's intestine preparation. The weaker contraction to 0·1μg. histamine, however, is not diminished by this concentration of cocaine; this shows that histamine acts on the muscle fibres only.
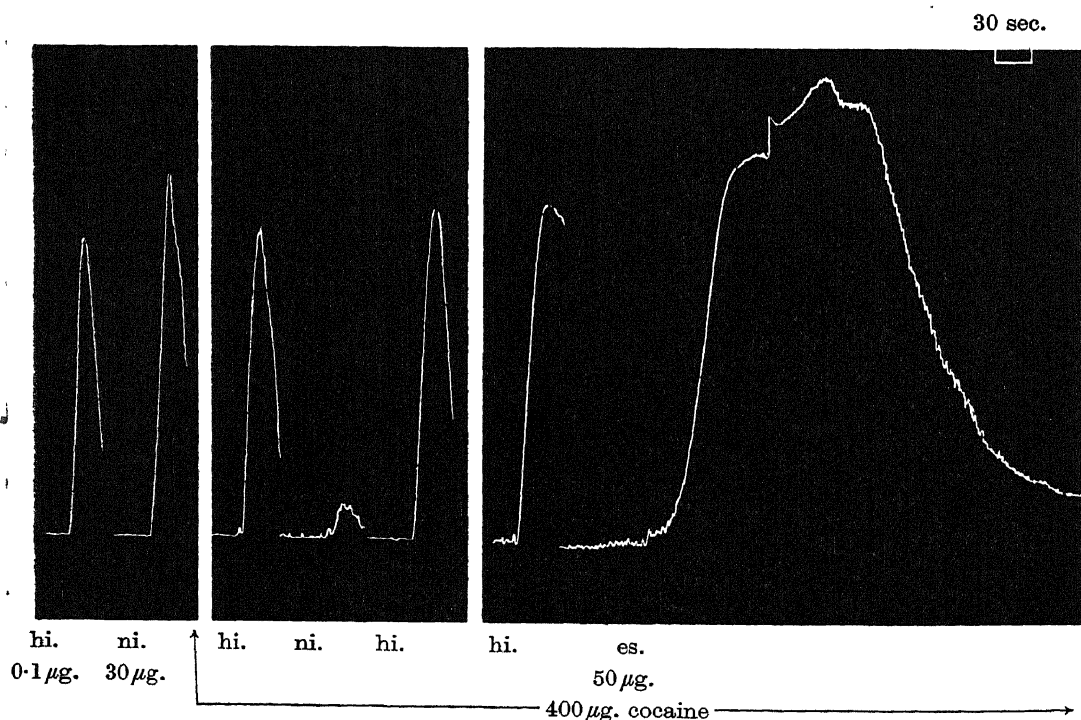
FIGURE 1. Contractions of guinea-pig's ileum preparation in 15 ml. bath to 0·1μg. histamine base (hi.), to 30μg. nicotine tartrate (ni.) and to 50μg. eserine sulphate (es.). Histamine and nicotine left in bath for 30 sec., eserine for 90 sec. After the first two contractions 400μg. cocaine chloride which was left in bath till end of experiment. (Time in 30 sec.)

On such an isolated intestinal preparation the spontaneous release of acetylcholine can be demonstrated by adding eserine to the bath. The released acetylcholine is then no longer destroyed, accumulates, and contracts the muscle. The eserine response persists when the nervous structures are inactivated by cocaine, as illustrated in the second part of figure 1. There is thus good evidence that the spontaneous release of acetylcholine from the intestinal wall persists after inactivation of the nervous elements.

Distribution of choline acetylase in the wall of the digestive tract

If acetylcholine is continuously released it must be replaced and resynthesized. There is thus the problem: is the enzyme responsible for the synthesis of acetylcholine, the choline acetylase located in those parts which contain the nerve cells?

The wall of the small intestine consists of several layers, and in large animals, as in dogs, five layers can easily be separated. Starting from the lumen we have the mucosa, which can be separated into glandularis mucosa and muscularis mucosa. The term glandularis mucosa refers to that part which is easily scraped off from the muscularis mucosa with a knife, and which consists mainly of the glandular tissue. The muscularis mucosa is a thick sheath of smooth muscle fibres. The mucosa is separated from the outer muscle layer, the external muscularis, by a tough membranous sheath of connective tissue, the submucosa, which is easily removed *in toto*. The external muscularis consists of a thick layer of circularly arranged muscle fibres adjacent to the submucosa and a thin layer of longitudinally arranged muscle fibres covered by the serosa.

In guinea-pigs the cells of the submucous plexus may extend to the glandularis mucosa. In men and dogs they are located in the submucosa; no histological examination, however, was made to find out if the nerve cells adhere mainly to it or to the muscularis mucosa when these layers are separated from each other. But the submucosa was the only layer in which no choline acetylase could be detected, whereas all other layers of the wall of the small intestine contained the enzyme.

The myenteric plexus is situated between the circular and longitudinal muscle layers, and the nerve cells adhere to the latter when it is stripped off. Therefore, if the enzyme were located in the nerve cells, only the longitudinal muscle layer should contain it. This was not found.

Table 2 gives the amounts of acetylcholine synthesized per gram dried tissue for duodenum (Du.) and ileum (Il.). In the first two columns the amounts are given per gram layer. The nerve-cell free glandularis mucosa gives the highest values, and there are no significant differences between longitudinal and circular muscle. In order to obtain an exact picture, however, of the distribution of the enzyme in the various layers, these values are not sufficient. The thickness of the different layers has to be taken into account as well. Columns 3 and 4 show the differences in the thickness of the different layers and, in the last four columns, the synthesis of acetylcholine for each layer is given per gram wall or as percentage of that of the whole wall. It will be seen that the thin longitudinal muscle layer, which contains as much, or even more, choline acetylase per gram tissue than the circular muscle layer, contributes much less to the total enzyme content of the whole wall than the

thick circular muscle layer. Thus again, the percentage distribution of the enzyme in the wall bears no relation to the localization of the nerve cells.

TABLE 2. DISTRIBUTION OF CHOLINE ACETYLASE IN LAYERS OF DOG'S SMALL INTESTINE (FROM FELDBERG & LIN 1950)

layers	μg. acetylcholine per g. *layer*		mg. tissue per g. wall		μg. acetylcholine per g. *wall*		percentage distribution of enzyme in wall	
	Du.	Il.	Du.	Il.	Du.	Il.	Du.	Il.
gl. mucosa	112	98	250	223	28	22	39	32
m. mucosa	53	57	425	283	23	16	31	23
mucosa			675	506	51	38	70	55
circul. mus.	61	66	226	374	14	25	19	36
longit. mus.	76	50	99	120	8	6	11	9
mus. ext.			325	494	22	31	30	45
whole wall	73	69	1000	1000	73	69	100	100

TABLE 3. ACETYLCHOLINE SYNTHESIS AND NUMBER OF NERVE CELLS PER SQ.CM. OF MUSCULARIS EXTERNA IN GUINEA-PIG (FROM FELDBERG & LIN 1950)

	acetylcholine (μg./hr./cm.^{2})	nerve cells per cm.^{2}	ratios	
			synthesis	nerve cells
oesophagus	0·000	1,350	0	0·4
stomach (cardia)	0·032	3,500	1·0	1·0
stomach (fundus)	0·065	7,000	2·0	2·0
stomach (pyloric)	0·305	18,000	9·5	5·1
duodenum	0·352	10,000	11·0	2·9
ileum	0·440	7,500	13·8	2·1
caecum	0·155	4,500	4·8	1·3
colon	0·304	18,000	9·5	5·1

There is one further point. Even if the nerve cells are not the site of the acetylcholine synthesis, this does not exclude nervous elements altogether, because nerve fibres extend to all layers. But a certain relation between cells and fibre density is to be expected.

For the guinea-pig's digestive tract figures are available about the density of nerve cells in the muscularis externa. Table 3 compares the concentration of the choline acetylase with the density of the nerve cells in the different parts of the digestive tract. The figures for the nerve cells are taken from a paper by Irwin (1931). The concentration of both tissue constituents changes in the same direction in the different sections, but on a more quantitative basis this parallelism breaks down. A doubling of the number of nerve cells may be associated with a nearly fourteenfold increase in synthesizing power.

Functions of the released acetylcholine

The inability of cocaine to stop the release of acetylcholine as well as the results on the distribution of choline acetylase suggest that the acetylcholine metabolism in the wall of the intestine is, at least to a large extent, non-nervous in origin,

comparable to the acetylcholine metabolism of the non-nervous human placenta and probably associated with the choline and lipin metabolism in these organs.

For attributing a physiological function to this phenomenon we rely not solely on the finding of choline acetylase and of acetylcholine in the various layers, but also on the responses which have been obtained (Hambleton 1914; King & Arnold 1922; Verzar & v. Kokas 1927; Beznak 1936; Wright, Jennings, Florey & Lium 1940; Holton 1949) when acetylcholine and eserine are allowed to act on these layers and how these responses are affected by atropine.

By its action on the smooth muscle fibres the released acetylcholine will increase muscle tone and motility and thereby facilitate the peristaltic reflex. This motor action is not confined to the muscularis externa but applies to the muscularis mucosa as well. By its action on the gland cells in the mucosa, on the other hand, the released acetylcholine provides a physiological stimulus for the secretion of succus entericus. The release of acetylcholine in the mucosa may also influence absorption.

There is a further possibility of action. The released acetylcholine may stimulate not only the gland cells and muscle fibres in the intestinal wall, but the nerve cells as well. On this assumption they would seem to be imbedded in a medium which received continuously their stimulating agent, acetylcholine. This conception is a reversal of what was thought a few years ago. Then the nerve cells were considered to be the source of the acetylcholine synthesis. Now we postulate a non-nervous origin for the synthesis and consider the possibility that the released acetylcholine may then act, in its turn, on these nerve cells to keep them, so to speak, in a continuous state of tonic activity. There is some evidence for this assumption.

Stimulation of the cells of the myenteric plexus by acetylcholine

Some time ago Ambache (1946, 1949) published experiments which were meant to show that the contractions of the intestinal preparation, in response to small doses of acetylcholine, are mainly due to stimulation of the cells of the myenteric plexus. After large, paralyzing doses of nicotine, small doses of acetylcholine, which previously had caused strong contractions of the preparation, became inactive, and the original, strong responses could not be obtained even when the doses of acetylcholine were increased tenfold. The reduction in the acetylcholine response was attributed to paralysis of the ganglion cells, but this conclusion is not justified because the excitability of the muscle fibres is reduced as well (Emmelin & Feldberg 1947). For instance, responses to histamine are reduced as much as, and sometimes even more than, those to acetylcholine (see figure 2). Histamine, however, has no stimulating action on autonomic ganglia.

But we have other inhibitors of ganglionic activity; D-tubocurarine, for instance, paralyzes the ganglion cells in the intestinal wall. As seen from the experiment of figure 3, the addition of 500 μg. D-tubocurarine to the 15 ml. bath abolished the stimulating action of 40 μg. nicotine. This dose of D-tubocurarine augmented the response to 0·2 μg. histamine, but reduced that to 0·15 μg. acetylcholine. One might easily be tempted to attribute this reduction to paralysis of the ganglion cells.

30 sec.

ac.　hi.　ac.　hi.　ni.　　hi.　ac.　　hi.　　ac.　hi.　ac.　hi.

100 μg.
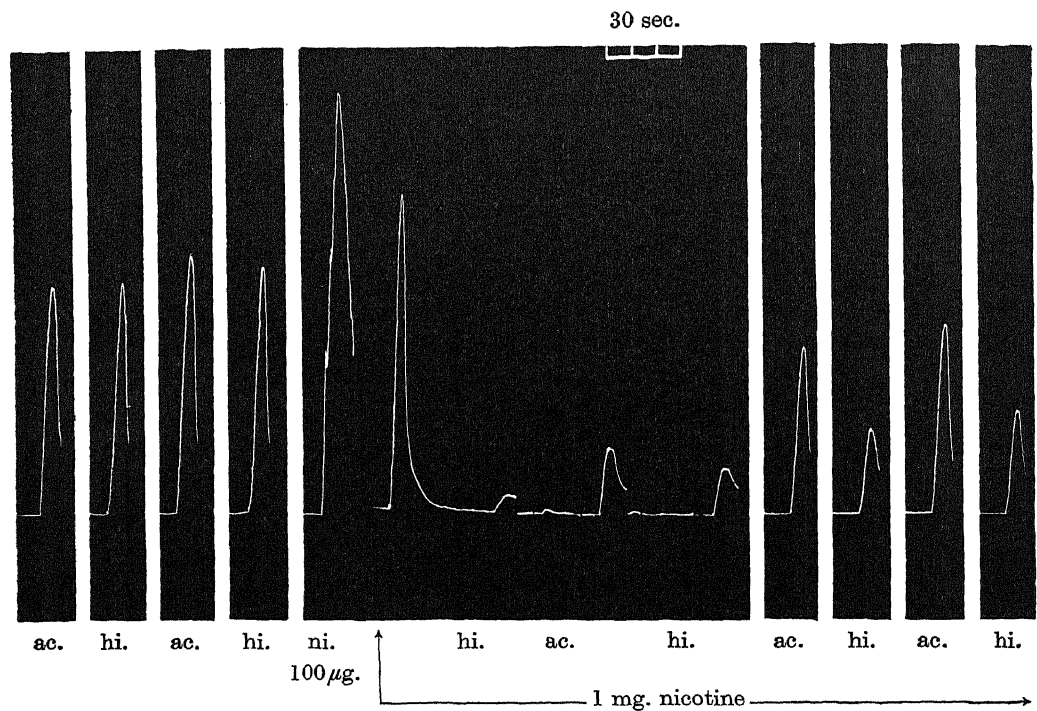
1 mg. nicotine

FIGURE 2. Contractions of guinea-pig's ileum preparation in 15 ml. bath to 0·3 μg. histamine base (hi.), to 0·25 μg. acetylcholine chloride (ac.) and to 100 μg. and 1 mg. nicotine tartrate (ni.) respectively. All drugs left for 30 sec. in bath except the large dose of nicotine, which was left till end of experiment. (Time in 30 sec.)

30 sec.

ac.　ac.　ac.　ac.　　hi.　ni.　　hi.　ni.　　hi.　hi.　hi.　hi.　hi.　hi.　hi.　hi.

0·15 μg.　　　500 μg.　　0·2 μg. 40 μg.　　500 μg.　　　　2000 μg.

D-t.c.　　　　　　D-t.c.　　　　　　D-t.c.
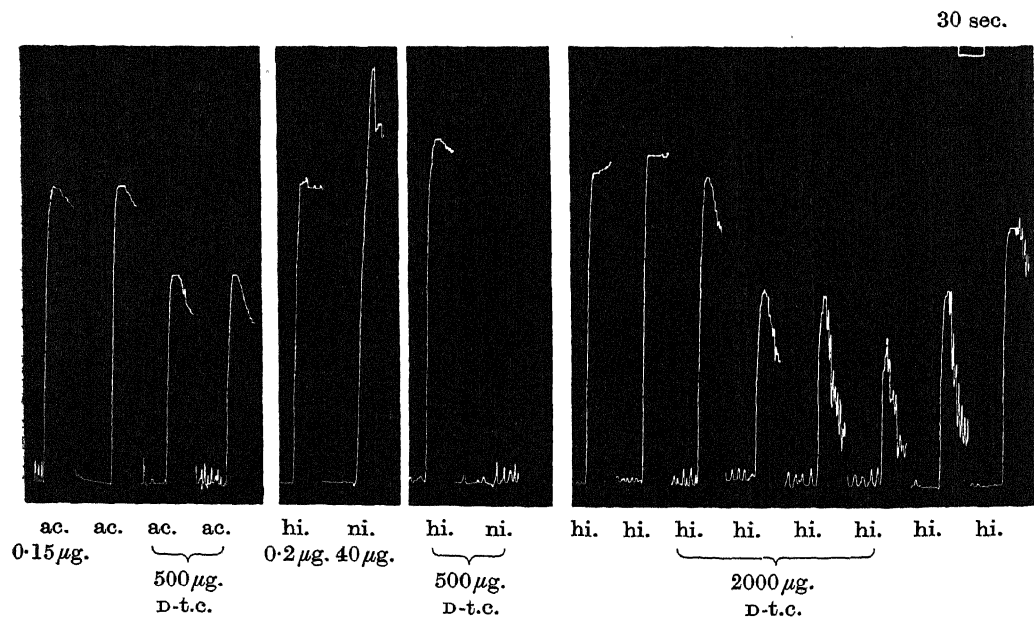
FIGURE 3. Contractions of guinea-pig's ileum preparation in 15 ml. bath to 0·15 μg. acetylcholine (ac.), 0·2 μg. histamine base (hi.) and to 40 μg. nicotine tartrate (ni.) in the presence of 0·5 mg. and 2 mg. D-tubocurarine respectively, as indicated. Acetylcholine, histamine and nicotine left in bath for 30 sec. (Time in 30 sec.).

The problem is not as simple as this. If we increase the dose of D-tubocurarine two- or fourfold, the response to histamine becomes reduced as well, as seen in figure 3. In this respect D-tubocurarine resembles atropine, and the reduction of the acetylcholine response might, therefore, well be the result of an atropine-like action of D-tubocurarine. The possibility that some ganglionic inhibitors have atropine-like actions would explain why different ganglionic inhibitors, which all abolish the nicotine response, affect the acetylcholine response to a different degree. Tetraethylammonium, for instance, has a weaker effect on the acetylcholine response than D-tubocurarine when tested on the same preparation. As Collins (1948) has shown, it does not even regularly reduce the acetylcholine response.

In using inhibitors of ganglionic activity on the intestinal preparation we need substances which have no atropine-like effects; that is, substances which do not reduce the histamine response even when used in concentrations stronger than those required to abolish a nicotine response.

Recently, Paton & Zaimis (1949) have shown that hexamethonium is a strong inhibitor of autonomic ganglia. As an inhibitor of intestinal ganglia, it is more than twice as strong as D-tubocurarine. Hexamethonium increases and never reduces the histamine response; apparently it has no atropine-like action. It is of interest to note, in this connexion, that all inhibitors of the nervous structures in the intestinal wall so far examined—D-tubocurarine, tetraethylammonium, hexamethonium and cocaine—in certain concentrations sensitize the intestinal preparation to the action of histamine.

When hexamethonium is tested on the acetylcholine contractions of the guinea-pig's gut, it does not produce a reduction as strong as that produced by D-tubocurarine, although it abolishes the nicotine response. Sometimes the acetylcholine response is reduced slightly; sometimes it is not reduced at all; it may even be augmented. But even then there is this difference: the histamine response is usually augmented more than the acetylcholine response. It looks as if a small portion, often less then 5 % but sometimes up to 20 % of the acetylcholine contractions on the guinea-pig's small intestine, results from ganglionic stimulation. To a certain extent Ambache's original suggestion may therefore well be correct, and there is some justification for assuming that the acetylcholine continuously released in the intestinal wall either stimulates the nerve cells, or at least lowers their threshold to other stimuli.

If such a mechanism occurs in the intestinal wall, it may well be argued: why do we not assume a similar mechanism in the central nervous system, even if it were difficult to see where the acetylcholine should come from in this instance, if not from the nervous elements? Nevertheless, it is surprising how many facts could be reconciled with this assumption. For instance, that acetylcholine stimulates nerve cells, that eserine increases central nervous activity, and even the fact that in many parts where the nerve cells are situated, the grey matter, there is a high concentration of choline acetylase. But there are other facts which are difficult to reconcile with this hypothesis. There are many white tracts in the central nervous system which contain the enzyme in high concentration and, therefore, must be cholinergic neurons; for instance, the lateral filet and trapezoid bodies, the fimbria and

fornix and the thalamic mammillary tracts (Feldberg, Harris & Lin 1950). There is, in fact, not sufficient evidence to assume such a hypothesis for the central nervous system. But as an instance to illustrate how the new conception of Professor Burn concerning the functions of the chemical transmitters may influence our conventional ideas, it is worth while to mention this rather disturbing hypothesis.

THE ACTION OF LOCAL HORMONES

By J. H. GADDUM, F.R.S.

In order to account for the resemblance between the actions of sympathetic nerves and those of adrenaline, Elliott (1904) suggested that sympathetic nerves might act by liberating adrenaline. Discussing this theory Barger & Dale (1910) came to the conclusion that the actions of sympathetic nerves did not resemble those of adrenaline so closely as they did those of the corresponding primary amine now known as noradrenaline. Loewi's demonstration that some sympathetic nerves do actually liberate a substance like adrenaline was published in 1921.

Cannon & Rosenblueth (1937) obtained evidence that stimulation of the sympathetic nerves in a cat led to the liberation into the blood stream of a variable mixture of two substances. They supposed that one of those substances had purely excitor effects and the other purely inhibitor effects, and they called them sympathin E and sympathin I. The discovery that two substances were involved was an important advance, but it now seems likely that the two substances are adrenaline and noradrenaline, and since neither of these has purely excitor or purely inhibitor effects the terms sympathin E and sympathin I should be forgotten. The word sympathin is a convenient label for the substances specifically liberated by adrenergic nerves and may perhaps survive with this meaning. The reason for the slow advance of knowledge about these substances is that the quantities available for study are small. The evidence is still almost entirely pharmacological, since no other methods are sensitive enough. The distinction between adrenaline and noradrenaline has depended on the use of parallel quantitative tests using tissues with different values for the ratio of the activities of these two drugs. If the unknown substance is adrenaline it should match the same amount of adrenaline by any test. It was by using this criterion and studying the effect on distant organs of substances liberated into the general circulation that Cannon & Rosenblueth showed that adrenaline was not the only substance liberated. Their results were such as might be produced by the liberation of a mixture of adrenaline and noradrenaline, but they did not adopt this explanation which was first published by Bacq (1933).

Attempts were made to test this theory by methods similar to those of Cannon & Rosenblueth, but the results were not conclusive (Greer, Pinkston, Baxter & Brannon 1938; Gaddum & Goodwin 1947). When substances are liberated at an unknown rate into the natural circulation the result is affected by too many variables. Experiments on fluids collected in test-tubes provide better evidence, but are difficult owing to the small amount of the active substances which can be collected in this way.

In recent years new evidence has accumulated in various ways. Euler (1946, 1948) and Bacq and their colleagues have obtained evidence that extracts of various tissues contain noradrenaline. The concentration is particularly high in adrenergic nerve trunks. A comparison of colorimetric estimates with biological estimates of the amount present in an extract of splenic nerves showed that the substance was present as the laevo-isomer in a concentration of 10 to 15 μg./g. With such high concentrations the experiments are comparatively easy and the result reasonably certain. Confidence in these findings was, nevertheless, increased by other evidence that noradrenaline was a natural substance. Holton (1949) showed that adrenal medullary tumours contained enough noradrenaline to be identified as a separate spot on a paper chromatogram. Evidence of the presence of noradrenaline in adrenal extracts was obtained almost simultaneously in America, Sweden and Germany, and it was actually isolated from this source (Tullar 1949).

These results have established the natural occurrence of noradrenaline, but have no direct bearing on local hormones. The evidence regarding the substances liberated by nerves may appear at first sight to be conflicting. Early work in various laboratories with isolated organs seemed to show that adrenaline was liberated in the frog's heart (Loewi 1936), the rabbit's ear (Gaddum & Kwiatkowski 1938) and in a sympathetic ganglion (Bülbring 1944). In view of the evidence that other substances might be involved Peart (1949) undertook an elaborate study of the substances liberated by the cat's splenic nerves. He showed that the main substance was noradrenaline, but this was mixed with 10 % or less of adrenaline. This has led to reconsideration of some of the previous evidence. Outschoorn has done more experiments with perfused rabbits' ears in my department. His results support the view that *both* adrenaline *and* noradrenaline are liberated when the sympathetic nerves are stimulated in this preparation. The evidence which led Bülbring to identify the substance liberated in the superior cervical ganglion as adrenaline depended on quantitative agreement between tests on perfused frog's heart and pigeon's rectum, and colorimetric and fluorimetric tests. It is now known that in all these tests the dosage ratio (noradrenaline/adrenaline) for equivalent effects is in the range of 30 to 50. The evidence, therefore, does not exclude noradrenaline. In any case, it now seems probable that most adrenergic nerves liberate a mixture of adrenaline and noradrenaline. The factors controlling the percentage of methylation are still largely unknown. Progress is arrested while techniques are being developed for estimating small quantities of adrenaline and noradrenaline in mixtures of these substances. It is tempting to suppose that there is some relation between the facts that the difference between adrenaline and noradrenaline is a methyl group and that acetylcholine contains transferable methyl groups, but there is no evidence about this yet.

Consider now the differences in the effects produced by adrenaline and noradrenaline. Barger & Dale (1910) concluded that the essential difference consisted in the fact that noradrenaline failed to reproduce those actions of adrenaline which involved the inhibition of activity. These inhibitor actions also differed from excitor actions in the fact that it was difficult or impossible to inhibit them with ergotoxine. Konzett (1940) studied the effects of a series of substances with marked inhibitor

actions and feeble excitor actions. The most active member of this group, isopropyl-noradrenaline, now officially christened isoprenaline, has properties which are the opposite of those of noradrenaline; its actions are predominantly inhibitor.

Barger & Dale thus introduced the distinction between two types of action, one of which is produced by noradrenaline and inhibited by ergotoxine, while the other is not readily affected by either of these drugs. This distinction has played an important part in pharmacology, but later work makes it necessary to modify the original theory. The difference between the actions of adrenaline and those of noradrenaline are quantitative rather than qualitative. Practically all the actions of adrenaline are also given by noradrenaline in sufficient doses. In comparing their effects it is therefore convenient to measure the ratio of the dose of noradrenaline to the dose of adrenaline having the same action. According to Barger & Dale's theory this dosage ratio should be high for inhibitory actions and low for excitor actions. In a search for methods of identifying substances liberated by nerves, West (1947) estimated this dosage ratio for a number of different tissues. Its exact value depends on the conditions. West found, for example, that when pieces of isolated intestine were stored in the cold for several days and then warmed up and tested, their sensitivity to adrenaline had diminished, but their sensitivity to noradrenaline was well maintained. With fresh intestine the ratio was 1 to 3, but after several days it fell to about 0·2. When a frog's heart was isolated by Straub's method the dosage ratio was 8, but it became 33 when the heart was perfused, so that the drugs were only in contact with the tissues for a short time. West thought that this was because noradrenaline acted more slowly than adrenaline and obtained other evidence that this was so. However, even when full allowance was made for such factors it was clear that the ratio was low for intestine and high for the frog's heart. This was unexpected, since the action on the intestine is inhibitor and should have a high dosage ratio like the uterus, and the action on the heart is excitor and should have a low ratio.

Ahlquist (1948) compared the actions of adrenaline with those of four allied substances on a series of tissues. In each experiment he placed the drugs in the order of activity and so found that one group of tissues gave one order and another group of tissues gave almost exactly the opposite order. One group of effects was predominantly excitor and the other predominantly inhibitor, but Ahlquist agreed

TABLE 4

OH
|
HO—⟨benzene ring⟩—CHOH—CH—NH
                        | |

			α	β
noradrenaline	H	H	2	5
corbasil	CH_3	H	3	4
adrenaline	H	CH_3	1	3
N-methyl corbasil	CH_3	CH_3	4	2
isoprenaline	H	$CH(CH_3)_2$	5	1

α-effects. Contraction of arteries (skin), uterus (rabbit or pregnant cat), nictitating membrane, or ureter. Inhibition of intestine.

β-effects. Inhibition of arteries (voluntary muscles and heart), uterus (rat or virgin cat) or bronchi. Acceleration of heart.

with West that the intestine and the heart were anomalous. In discussing his results he speaks of α-receptors and β-receptors, but it is sufficient for the present purpose to speak of α-effects and β-effects.

The α-effects of this classification include the excitor effects of Barger & Dale and the inhibitor effect on the intestine; they are readily produced by noradrenaline. The β-effects include the inhibitor effect on the blood vessels, and on certain kinds of uterus and quickening of heart. Noradrenaline has these β-effects only in comparatively high doses.

The actions of ergotoxine and other antagonistic drugs fit in well with this classification. When ergotoxine was thought to antagonize excitor effects only, it seemed odd that it antagonized the inhibitor effect of adrenaline on the intestine even in quite low concentrations (Clark 1937), but had little action against the excitor effect on the heart. With the new classification these anomalies disappear; ergotoxine antagonizes α-effects more readily than β-effects.

Isoprenaline has powerful β-effects and comparatively feeble α-effects. Its dilator action on the blood vessels and the bronchi and its accelerator action on the heart are greater than those of adrenaline, but it has a comparatively feeble action on the intestine (Konzett 1940; Ahlquist 1948).

Hyperglycaemia is difficult to classify, as the evidence is inadequate and conflicting. Adrenaline appears to have more effect than either noradrenaline or isoprenaline. The fact that noradrenaline has much less effect than adrenaline on the ascorbic acid content of the adrenals (Nasmyth 1950) suggests that this is a β-action. The best way of deciding how to classify an effect is probably to compare the action of noradrenaline with that of isoprenaline. The data are not available in every case. It is difficult to explain all these facts except on the theory of two types of receptor, but perhaps it is permissible to forget them for a while and speculate on fundamental mechanisms.

It was shown by Clark (1937) that the curves connecting the dose of local hormones and other similar drugs and their effect upon plain muscle were similar in shape to the dissociation curve of oxyhaemoglobin. This led to a simple picture of what happens at a receptor. A molecule of local hormone combines with a receptor, remains in combination for a finite time and then diffuses back into the surrounding fluids. However, any other mechanism which removes the molecules from receptors at a rate proportional to the number combined would fit the data equally well. It is possible, though not essential, that the receptor forms part of an enzyme which plays an essential part in the metabolism of the cell and that the local hormone is a substrate for this enzyme.

The action of the drug is assumed on this theory to be proportional either to the amount of drug combined with the enzyme at any given moment, or to the rate of action of the enzyme on the drug. Since these two quantities are likely to be proportional to one another it makes no difference which is considered. When the enzyme has acted upon the drug the products of the reaction may perhaps be acted upon by another enzyme, or a series of enzymes. If this theory is correct it might be expected that drugs which antagonize the actions of local hormones would act by paralyzing one of the enzymes involved. The actions of these antagonists on

enzymes have not yet been studied in sufficient detail to prove or disprove this theory. Quantitative studies suggest competitive antagonism between local hormones and their antagonists, but do not suggest what these drugs are competing for, and very little is known about this. Is it possible that acetylcholine combines with the enzyme system which is responsible for maintaining the high concentration of potassium in cells?

In considering the mode of action of adrenaline antagonists it should not be forgotten that some at least of them may have the opposite effect in small doses. Jang (1940) collected various examples of this phenomenon from the literature and showed that low concentrations of ergotoxine, yohimbine or F 933 (piperidyl-methylbenzodioxane) sensitized the vessels of the perfused ear of a rabbit to adrenaline. In higher concentrations these drugs were antagonistic to adrenaline. This work arose out of similar work on ephedrine and other sympathomimetic amines which have the same double action, causing sensitization in small doses and antagonism in large doses. Gaddum & Kwiatkowski (1938) upheld the view that the sensitizing effect of small doses of ephedrine was due to inhibition of an enzyme which destroys adrenaline. In recent years we have been reproached for thinking that all the actions of ephedrine were due to this effect. In actual fact, we discussed this theory in our original paper and gave reasons for rejecting it. Some of the actions of ephedrine could certainly not be explained in this way. What we did actually say has also been criticized; but for the wrong reasons. The theory is not disproved because Richter & Tingey (1939) failed to find amine oxidase in the rabbit's ear, since Professor R. H. S. Thompson has been more successful and has shown that the enzyme is present in the artery as the theory requires. The theory is not disproved by the fact that sensitization is produced by very low concentrations of ephedrine, since it is possible that the tissue concentrates the drug at the appropriate site. The main objection to the theory lies in the fact that ephedrine sensitizes the tissue not only to adrenaline but also to corbasil, and corbasil is not destroyed by amine-oxidase (Jang 1940). One way to avoid this objection is to say that amine-oxidase is not the enzyme involved. The theory explains facts not otherwise explained and should not be rejected without good reasons.

As Professor Burn has said, the direction of the effects of local hormones themselves may also depend upon the dose; excitation and inhibition may occur as two phases of the response to a single dose. The explanation of these opposed effects is obscure. The inhibitory effect of large doses might perhaps be comparable with the inhibitory effect of high concentrations of substrate upon certain enzymes (Zeller, Schär & Staehlin 1939; Lardy 1949). This may be due to the accumulation of the products of the reaction. The insensitivity which may develop when tissues are exposed for a few minutes to drugs might be comparable with the adaptation which occurs in sense organs, or with the accommodation of excitable tissues to electric potentials. Such theories are unlikely to provoke discussion, but consider the electric response.

There is much evidence that when local hormones are applied to tissues they produce electric changes which are presumably due to the depolarization of membranes. The first effect is likely to be instability of the membrane, which rapidly

alternates between polarization and depolarization. The second effect of large doses is likely to be a more or less continuous state of depolarization. This appears to be what happens when acetylcholine acts upon an autonomic ganglion or a voluntary muscle; an initial burst of conducted impulses is liable to be followed by a state of inhibition in which the tissue is insensitive to the drug. In this case it is clear that the unstable state causes activity and depolarization causes inhibition. It is un-necessary in this connexion to attach any precise meaning to the words polarization and depolarization. They are used as labels for two states between which it is assumed that the surface may alternate.

The evidence of the electrophysiologists about what happens when local hormones act on plain muscle is more confusing (Fischer 1944). The tissues studied generally contain nerve fibres and synapses which complicate the picture. Unstable states and frequent electric changes appear to be associated with contractions. It is commonly assumed, however, that a continuous state of depolarization is also associated with contractions, or rather with contractures. Some of the pharmacological data could be explained on the theory that a more or less continuous state of depolarization may be associated in some tissues with inactivity and relaxation. This may seem inconsistent with results such as those of Bacq & Monnier (1935), who recorded slow monophasic electric changes corresponding in time to the mechanical response. Such records may, however, be due to the summation of brief electric changes in numerous muscle fibres.

The response of the smooth muscle of the gut or uterus to histamine appears to resemble the response of the ganglion to acetylcholine. The contraction of the muscle due to a large dose is followed by an inhibitory state. Can this be another example of an active unstable state, followed by depolarization and inactivity? There seems to be no direct evidence at present.

By R. A. PETERS, F.R.S.

This discussion is devoted to local hormone action, i.e. 'hormones' inside cells. In the last limit I do not feel that it is easy to separate this from the action of such important constituents of the cell as co-enzymes; ought we for instance to regard vitamin B_1 pyrophosphate as some species of local hormone; certainly its absence can alter profoundly the action of the central nervous system. I think, however, that we are meant to exclude the normal enzymic machinery, and to confine ourselves more to those substances which may speed or slow enzyme processes in tissues. Roughly, outside the substances which have been discussed already, this amounts to saying: do we know of internal cell metabolites which can modify the activity of organized enzyme systems so that physiological changes or pharma-cological responses can be produced?

Confining ourselves to this we can soon see that there are theoretical possibilities. For instance, it is now well known that several enzymes require for their functioning an active —SH group; and that if this disappears the enzymes are blocked with a corresponding interference with metabolism; iodoacetate blocks especially the

triosephosphate dehydrogenase, and the trivalent arsenicals, enzymic components of the pyruvate oxidase system. If the —SH groups of enzymes are important for their functioning, then the maintenance of active —SH groups must matter, and it is reasonable to think (as suggested by Hopkins's work years ago) that glutathione may play a part in regulating enzymic activity. Such actions would be rather direct.

There is, however, a somewhat more subtle way in which organized enzymes in tissue complexes may be thrown out of action, to which I wish to draw especial attention, as it seems directly pertinent to this discussion. To come straight to the point, following upon some work with fluoroacetate *in vitro* in kidney homogenates with C. Liébecq (Liébecq & Peters 1949), Dr Buffa and I have recently published estimations in various organs of the rat and other animals of citrate, following injections of fluoroacetate (Buffa & Peters 1949). This is a poison which induces either strychnine-like convulsions of the nervous system or exerts toxic effects upon the heart. Within an hour, very high accumulations of citrate were found in several organs, some of which are shown in table 5. In accounting for these, two things must be remembered: (1) Owing to the stability of the C-F bond, there is no possibility of this compound blocking SH groups like the arsenicals or like iodoacetate. (2) The compound does not block any individual isolated enzyme—at least, many have been tried without success. The accumulations of citrate receive a ready explanation if we believe that the 'tricarboxylic acid' cycle is operating *in vivo*.

TABLE 5. ACCUMULATION OF CITRATE IN THE RAT (BUFFA & PETERS 1949)

	citric acid μg./g. wet tissue	
	control	poisoned*
brain	21	166
heart	25	677
stomach	37	386
small intestine	36	368
diaphragm	0	400

* NaFlAc 5 mg./kg. injected intraperitoneally 60 min. before animals sacrificed for the estimations.

This system of enzymes is now considered by biochemists to exist in the mitochondria, and is included in what in my laboratory we have previously called the pyruvate oxidase system. It is the important and useful generalization of Krebs, according to which pyruvate is oxidized by undergoing a preliminary condensation as a 2-carbon fragment with a 4-carbon acid to form a 6-carbon three carboxylic acid, which is subsequently degraded to a 4-carbon acid with loss of the pyruvate as CO_2 and H_2O.

If we take the view that the tricarboxylic acid cycle is operating *in vivo*, and that the fluoroacetate can be activated like acetate and then built in to form a fluoro 6-carbon acid, there is at once a possible explanation of the citrate accumulations, as such a fluoro 6-carbon acid may well be too 'foreign' for further metabolism and may therefore 'jam' the 'cycle'. The hypothesis proposed is

represented diagrammatically in figure 4. We do not yet know which is the 6-carbon acid first formed, whether citric itself or whether it is the acid with the double bond known as *cis*-aconitate; but all these acids are in equilibrium with one another through the enzyme aconitase. The diagram which we have given embodies part of a previous hypothesis of Bartlett & Barron—who think that fluoroacetate just competes with acetate and prevents oxidation of acetate—and with some work by Saunders (1947) and colleagues.
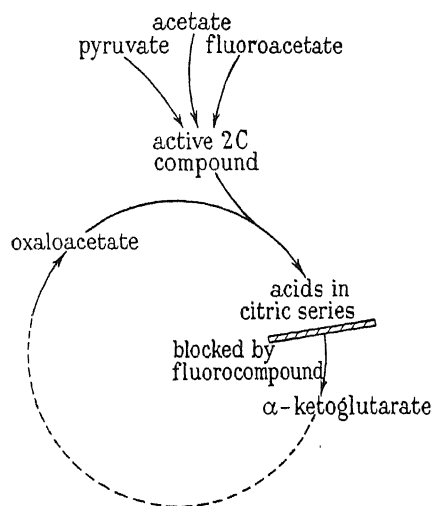
Figure 4. The diagram represents the course of reactions in the 'jamming' hypothesis of fluoroacetate action, as suggested by Liébecq & Peters (1949) and extended by Buffa & Peters (1949). It incorporates the idea of Bartlett & Barron (1947) that fluoroacetate competes with acetate in metabolism and is consistent with that of Saunders (1947).

We have reason to think that the actual block is at the stage of *iso*citric acid oxidation by the *iso*citric dehydrogenase, and that it is induced by the inhibitory action of a fluoro 6-carbon acid. It cannot be fluoroacetate itself;* Dr Lotspeich has confirmed in my laboratory that fluoroacetate has no effect on the *iso*citric dehydrogenase; hence it would have to be some compound formed from this. However this may be, the accumulation of citrate is a fact, which Buffa and I have established with R. W. Wakelin (unpublished experiments) by isolation of citric acid. As no other way is known in which citrate can be formed enzymatically, it seems to be highly probable that it is formed via the tricarboxylic acid cycle.

What we should like to put before you is the speculation that increases and decreases of citrate (formed enzymatically within the tissue cells) may be one of the possible regulators of ionic calcium activity. Brink, Bronk & Larrabee (1946) have shown the effects of small concentrations of sodium citrate ($380\,\mu g./ml.$) in increasing the excitability of nerve; the concentrations used and the effects produced are well in line with the convulsions taking place after injection of fluoroacetate.

* The 'jamming' hypothesis was suggested independently by C. Martius (1949), based upon *in vitro* experiments.

In order to make quite certain that these citrate accumulations can take place in frog's nerve, the following experiment has been made. A batch of frogs was injected with fluoroacetate and the citrate accumulation tested after 12 to 14 hr. against controls injected with saline. The animals got convulsions. Table 6 shows that there was a definite accumulation in the spinal nerves (brachial and sacral plexus) amounting to $80\,\mu g./g.$ tissue. This is only a preliminary experiment and the amounts are small; but when we consider that the nerve is not homogeneous, they seem to be definite enough to immobilize some of the calcium. What we therefore suggest as a speculation is that in the course of normal metabolism it is worth seeing whether normal activity may be controlled, at any rate in part, by the interaction of calcium (ionic) with citrate reached via the tricarboxylic cycle.

TABLE 6. INCREASE OF CITRATE IN THE NERVOUS SYSTEM OF FROGS, 12 TO 14 HR. AFTER THE INJECTION OF NA FLUOROACETATE (BUFFA & PETERS 1949)

	citric acid ($\mu g./g.$ wet tissue)	
	control	poisoned
spinal nerves	152	232
central nervous system	145	213

ON THE CONCEPT OF LOCAL HORMONES

By Z. M. BACQ

This discussion on local hormones comes at the right time; those who started the idea and suggested the term should be congratulated.

Some kind of order must be accepted in the classification of hormones, and one should first agree to a clear definition of what the hormonal action is. The presence of a definite substance in a tissue is not sufficient proof, even if in a certain group of animals this substance is considered as a true hormone. For instance, is the adrenaline which is elaborated in special cells in the ganglion of the ventral nerve cord of annelids a hormone; is it liberated in well-defined physiological conditions; does it act locally on the nervous cells? Nobody knows.

My collaborators (Fischer & Lecomte 1950) have recently shown that in the parotoid glands of various tropical toads, adrenaline is pure, unmixed with noradrenaline even in the denervated glands. It is extremely concentrated in the secretion (up to 4 %), but nobody has ever demonstrated that adrenaline in this peculiar location has any, even a local, hormonal function. I think that the true local hormones known at present are acetylcholine, adrenaline and noradrenaline; their action is not limited to the molecule itself, their metabolic products are also active. Other substances active locally (histamine, Lewis's H substance, leuco-taxine, etc.) are not strictly physiological; they are liberated when cells are injured, even slightly, by proteolytic enzymes. They are thus signs of a biochemical lesion

in R. A. Peters's sense. Finally, some local hormones are pure artefacts, for instance, the cardiac 'substances actives' of Demoor and collaborators.

I intend to comment briefly on the following points:

(1) Acetylcholine is a normal constituent of many tissues of both vertebrates and certainly invertebrates (Bacq 1947); its concentration is a characteristic chemical function of the tissues controlled by cholinergic nerves (MacIntosh 1938). In some invertebrates (annelids, holothurians, sipunculides) its function is quite similar to that in vertebrates. In the crustaceans, choline and its esters do not play any role in synaptic transmission (Bacq 1947), but protect the ganglion against strychnine poisoning (Bonnet 1938). This fact, strange as it is, indicates that choline and its esters might have some important metabolic function in the crustaceans.

In the vertebrates, choline present at the nerve endings after hydrolysis of the liberated acetylcholine may be responsible for the 'secondary refractory period' described by Fredericq (1938). When a maximum stimulus is applied to the peripheral end of the cut vagus nerve of the turtle, 2 to 8 sec. after a first effective inhibitory stimulus, a second stimulus is completely without action. Various experimental facts led Fredericq to propose the hypothesis that choline is responsible for this puzzling phenomenon.

(2) Sympathin is a mixture in variable proportions of adrenaline and noradrenaline: there is complete agreement on this point between English, Swedish and Belgian pharmacologists. Adrenochrome and other oxidized products of adrenaline are again interesting substances through their action on bleeding time and capillary resistance. The bleeding time is increased by sympathetic denervation; it is shortened by injection of adrenaline or adrenochrome in physiological small amounts ($0 \cdot 1$ to $10 \mu g$. for a rabbit). The haemostatic action of a single injection of adrenaline lasts for 6 hr., it is maximal only 7 min. after the injection. Adrenochrome acts without delay (Derouaux 1941). Being a quinone and hydrogen acceptor, it inhibits mitosis (Lettre 1943) and may be one of the normal chemical regulators of mitotic activity (Bacq 1949; Friedenwald & Buschke 1944). Adrenochrome also catalyzes different steps in carbohydrate metabolism (Green & Richter 1937). There are many reasons to believe that part of the adrenaline liberated locally, or in the circulation, is oxidized to adrenochrome and then to fluorescent substances (Bacq 1949). The peculiar pharmacology of the inhibitory actions of adrenaline also suggests that an oxidized molecule is responsible for some of its effects on smooth muscle. Thus it seems that the physiological functions of adrenochrome are as important as those of adrenaline (Bacq 1949). A similar picture has not been obtained with noradrenaline, because its oxidation to the corresponding quinone has not yet been chemically demonstrated.

At the last meeting of the Laurentian Hormone Conference (Franconia, September 1949), M. L. Tainter drew attention to the recently discovered fact that the nerve axon in the electron microscope appears as a series of fibrous structures with dense edges, called neuro-tubules, wherein the poliomyelitis virus particles migrate from the periphery to the centres at a speed of 2 mm./hr. (de Robertis & Schmitt 1949), a kind of ultramicroscopic model of the giant axon of the squid. This fact might explain (without necessarily giving support to Nachmansohn's views) how the nerve

fibre controls the concentration of local hormones in the cells it innervates. There might be a constant, although very slow, flow from the nerve into the region of its endings.

(3) I quite agree with J. H. Burn when he says that no proof has been given that histamine is a humoral transmitter. Abnormal stimulations, injuries, mechanical or chemical, liberate histamine. Rather mild stimuli to the skin liberate leucotaxine (or substances having the same actions); more leucotaxine, with necrosine and probably a series of active polypeptides, are set free in the process of local inflammation. It is a local process, probably unrelated to nervous activity, since it is not possible to show much difference in the reaction of denervated skin to these abnormal stimuli. There is little, if any, evidence that these substances are free and active in the normal tissues; they are liberated when the cell structure is destroyed or damaged. They should be called 'local reaction substances', not 'local hormones'; the term hormone conveys the idea of a physiological function.

(4) Some sixteen years ago, L. Brouha and I had a rather heated argument with Demoor and his collaborators (Brouha & Bacq 1936). The isolated *left* auricle of a mammal generally beats very seldom and irregularly; if one adds an extract of Keith-Flack node or of endocardium, the beating becomes vigorous and regular. This is not as Demoor thought because some local cardiac hormone (or hormones) regulate the heart beat; it is because the extract contains a mixture of choline, adrenaline, histamine, potassium ions, amino-acids, all kinds of very active substances. In the light of the recent work done in this country by McDowall (1946), J. H. Burn and his collaborators (Burn & Vane 1949), it might be interesting to take as a test not the two isolated auricles, but only the left one, as Demoor used to do. Many of the tracings published by Burn (sudden activity, sudden stop, and so on) are similar to those published by Demoor, who used the isolated left auricle.

This shows how dangerous it is to conclude from experiments with crude extracts that a local hormone is needed for physiological activity.

(5) I should like to ask Professor Burn a few questions and point out some facts which agree or disagree with his idea.

(a) One point is not clear in my mind; is acetylcholine (or any unstable choline ester) necessary? Is Professor Burn certain that in suitable concentration, choline would not do as well as acetylcholine? Choline was very effective in Bonnet's experiments on crustacean ganglia. Choline is probably the main neurochemical regulating substance of the reticulo-endothelial system (Chèvremont 1945, 1948), and acetylcholine is only half as active as choline in the process of histiocytic transformation *in vitro* (Thomas 1949).

(b) In text-books it is generally said that acetylcholine or vagus stimulation decrease the excitability of cardiac muscle. A long time ago H. Fredericq observed an increased excitability, a clear shortening of chronaxie in heart muscle after stimulation of the vagus or exhibition of acetylcholine (Fredericq 1924, 1925 a, b, c, 1926, 1931); it is not therefore so surprising that in certain conditions acetylcholine can initiate the rhythmic beating of arrested auricles.

(c) I tried to reconcile two well-known facts with Professor Burn's idea and I did not succeed. After section of the preganglionic fibres of the cervical sympathetic,

the acetylcholine content of the superior cervical ganglion in the cat falls rapidly to low levels (MacIntosh 1938); at that time, when the acetylcholine synthesis seems to be depressed, the ganglion cells become spontaneously active (Govaerts 1939; Coppée & Bacq 1938). Similarly, the striped muscle of mammals fibrillates a few days after section of the motor nerve. Thus there appears a spontaneous activity when the control of the cholinergic nerve is lost. Obviously we must know more about the acetylcholine synthesis in chronically denervated structures and see how the facts fit with those observed by Burn and his collaborators.

(d) Finally, I want to express my satisfaction to hear the pharmacologists of this country trying to explain their results in biochemical terms. In a recent discussion about anticholinesterases at the Belgian Academy of Medicine, I said that in my opinion this tendency to link pharmacological facts with biochemical action was logical, fruitful and unavoidable. I met strong opposition, and the discussion which started last April has only just ended (Bacq & Weekers 1949 *a*,*b*; Heymans 1949 *a*, *b*). I shall not forget to mention your opinion when I return to my country.

EXCITATORY AND INHIBITORY EFFECTS OF THE CHEMICAL MEDIATORS

BY G. L. BROWN, F.R.S.

In his opening remarks, Professor J. H. Burn has adumbrated an extension of the function of the chemical mediators, beyond their accepted position as transmitters of nervous activity, to a participation in the general metabolic processes of the tissues which they excite or inhibit. In the last 20 years our knowledge of the effects of acetylcholine and of adrenaline on skeletal muscle and the ganglia of the autonomic system has greatly increased. There has not, unfortunately, been an equal advance in our knowledge of the mechanism of the muscarine-like effects of acetylcholine or of the effects of adrenaline on smooth and cardiac muscle. I wish, therefore, to review some of the actions of these substances on striped muscle and ganglion cells and to consider whether the excitatory and inhibitory effects of the chemical mediators on other tissues can be explained in similar terms.

To take, first, the action of acetylcholine on striated muscle: the end-plate zone of muscle is specifically sensitive to acetylcholine, the artificial application of which generates the end-plate potential (Buchthal & Lindhard 1942; Kuffler 1943) and from this in turn, if it attains sufficient magnitude, arises the propagated excitation wave involving the whole muscle fibre in an all-or-nothing response. A persistence of acetylcholine at the end-plate, produced by the action of an anticholinesterase, gives rise to a prolonged depolarization and repetitive firing of the muscle fibre as long as the end-plate potential stays above threshold. There is some evidence that persistence at the mammalian end-plate of high concentrations of acetylcholine can produce a neuromuscular block (Brown, Burns & Feldberg 1948); experiments on mammalian denervated muscle, on avian muscle and on amphibian muscle show that this block is produced by a persisting depolarization which evokes not

propagated all-or-nothing contractions of the muscle, but a localized contracture which blocks not only transmission from nerve to muscle, but conduction along the muscle fibre (Brown 1937; Brown & Harvey 1938). The existence of such a block, as it were, beneath the end-plate, is very clearly shown by the experiments of Burns, Paton & Vianna Dias (1950) in their analysis of the actions of decamethonium iodide. We can recognize two associated, but separable effects of acetylcholine on skeletal muscle, depolarization of the end-plate producing conducted responses of the muscle, and a more intense depolarization, spreading far beyond the end-plate, producing neuromuscular block and conduction block in the length of the fibre. In sympathetic ganglia, under admittedly most abnormal conditions, a similar blocking effect of acetylcholine can be observed (Brown & Feldberg 1936).

To what extent can these known properties of acetylcholine explain its actions on other excitable tissues? In the instance of smooth muscle organs such as the stomach, or bladder, supplied with cholinergic motor fibres, the analogy is easy; there seems to be no reason to postulate any mechanism differing very greatly from that existing in skeletal muscle. In other words, acetylcholine would appear here to act by depolarizing some sensitive region of the muscle and initiating a contractile process. The analogy in some instances is very close; I might draw attention to McSwiney & Robson's (1929) observations on isolated innervated stomach muscle strips, in which they showed that single vagus volleys produced 'twitches' of gastric muscle, and that peripheral summation could occur much as it does under certain conditions in skeletal muscle.

I have already pointed out that acetylcholine, applied artificially, can produce neuromuscular block and conduction block in the muscle fibre in such tissues as avian or amphibian muscle. It is tempting to attribute the inhibitory effects of cholinergic nerves to the action at, or near, their endings of what I shall term for convenience 'excess' of acetylcholine. In the instance of a tissue normally endowed with rhythmic activity, the inhibitory effect would be exerted by a block of transmission of excitation arising at some distant pacemaker and spreading through the muscular syncitium (cf. Bozler 1941). One could give as examples of such an inhibition, the abrupt relaxation of the active stomach evoked by a stimulation of the vagus which, in the inactive organ, produces an almost equally abrupt contraction (McCrea, McSwiney & Stopford 1925; McSwiney & Wadge 1928). It might well be possible to explain on these lines the inhibition of the beating heart which results from vagus stimulation or application of acetylcholine. There is, unfortunately, remarkably little experimental evidence for testing this hypothesis, and what there is gives little support. The experiments of Kolm & Pick (1921), in which they showed that the frog's auricle, stopped by treatment with excess Ca^{++}, could be sent into contracture by acetylcholine and muscarine give a hint in the right direction, but the electrical evidence is, as far as it goes, most unfavourable to the idea. Some 60 years ago Gaskell (1887) measured the potential difference in the tortoise auricle between the middle of the auricle and the damaged end. He found that vagus stimulation caused an *increase* in the resting potential. There are obvious difficulties in the interpretation of the Gaskell effect; its very existence has, indeed, been doubted (cf. Einthoven & Rademaker 1917). Monnier & Dubuisson

(1934 a,b), however, are convinced of its occurrence, and Monnier (1936) has gone so far as to conclude that the action of vagus and sympathetic on the heart can be explained entirely on the basis of their effects on cardiac resting potentials—polarization by the vagus and depolarization by the sympathetic. Whether this is true or not remains to be seen, but it is perfectly clear that the *direction* of the effect produced by vagus stimulation is precisely the opposite of what one would expect were it acting in the way suggested above, by producing conduction block through a local depolarization in a conducting pathway.

Draper & Weidmann (1950) have, very recently, examined the effect of acetylcholine on the membrane potential of the mammalian Purkinje system, measured with an intracellular microelectrode. They have been kind enough to allow me to quote what is only a very preliminary conclusion, that 'at the moment they see no connexion at all between the electrical manifestations of acetylcholine at the motor end plate and the effects of acetylcholine on cardiac muscle'.

Of the electrical effects of inhibitory cholinergic fibres on smooth muscle systems there appears to be no recent information. Brücke & Oinuma, as long ago as 1910, studied the electrical changes in the retractor penis of the dog using the string galvanometer. Their records show that stimulation of the sacral roots produced inhibition of spontaneous rhythmic activity and a diminution of the potential of the sacral end of the muscle in respect of the undamaged preputial end. This again is unsatisfactory evidence, but it fits with the general conception of Bacq & Monnier (1935) that excitatory nerves produce a depolarization and inhibitory nerves a polarization of the tissue they supply. There is nothing, in fact, to suggest that the inhibitory cholinergic nerves exert their effect through the development of the persisting depolarization which characterizes the local block produced by 'excess' of acetylcholine in striated muscle.

We are forced, then, to the conclusion that, although the excitatory effects of acetylcholine on smooth muscle tissues may be brought about by an action on sensitive regions like the motor end-plate, such evidence as there is gives no support to the idea that the inhibitory effects of cholinergic nerves are due to the local accumulation of an excessive amount of the transmitter.

Let us now turn to the action of adrenaline. Its effects on skeletal muscle are, compared with acetylcholine, trivial, but they differ in another and most interesting way. The predominating effect of acetylcholine is upon the neuromuscular transmitting mechanism, the predominating effect of adrenaline is upon the contractile process; Burns, Bülbring and I (Brown *et al.* 1948) were unable, in fact, to detect any direct action of adrenaline upon neuromuscular transmission in isolated nerve muscle preparations. There seems, however, to be good evidence of an effect of adrenaline on ganglionic transmission (Bülbring 1944) and of a decurarizing action on muscle exerted by adrenaline when it is injected into the whole animal, an effect which we were tempted to attribute to the distant liberation into the blood stream of K^+ ions. Whatever the reason for the decurarizing action of adrenaline, it differs from the direct effect on contraction by being sensitive to dibenamine (Maddock, Rankin & Youmans 1948), whereas the direct muscular effect is not; we see immediately the possibility arising of two, apparently distinct, effects of

adrenaline upon skeletal muscle, one on excitability and another on the contractile mechanism.

The action of adrenaline on the contractile process has proved susceptible of further analysis. In the first place, the increase in muscle tension is associated with a retardation of the spread of the action potential and with an increase in the resting potential of the muscle (Brown, Goffart & Vianna Dias 1950). Goffart & Perry (1950) have now shown that these phenomena are further associated with a reduction of the K^+ exchange between the muscle and the fluid bathing it. The subsequent re-establishment of the resting condition of the muscle membranes is associated, as might be expected, with an increased loss of K^+ and a corresponding fall in resting potential.

This analysis of the effects of adrenaline on skeletal muscle gives us, unfortunately, remarkably little clue to its effects, excitatory or inhibitory, on other excitable tissues. Examination of potential changes in a tissue like the nictitating membrane, excited by adrenaline or sympathetic stimulation, gives unmistakable evidence that adrenaline can act by producing a local depolarization from which arise in turn propagated muscular responses associated with action potential spikes (Eccles & Magladery 1937a,b). Injected adrenaline again can produce a local non-propagated depolarization and a contracture which, like the depolarization produced by acetylcholine, prevents the propagation of excitation arising from incident nerve volleys. The analogy between these excitatory actions of adrenaline and those of acetylcholine on avian or amphibian skeletal muscle is complete, and we must obviously view adrenaline as an agent, like acetylcholine, capable of evoking local potentials of varying magnitude at the sensitive regions of the appropriate excitable structure.

There, however, the analogy ceases; no relation apparently exists between its action on the contractile process in mammalian muscle and its effects on cardiac muscle. Acceleration of the beat of the pacemaker, increased conduction velocity, and general depolarization of the muscle (cf. Gaskell 1887; Monnier & Dubuisson 1934a,b) are phenomena which are apparently the direct opposite of its actions on skeletal muscle.

An almost equal obscurity surrounds the inhibitory actions of adrenaline. As with acetylcholine, it is difficult to find satisfactory evidence that its inhibitory actions are due to accumulation, or excess of a substance which in lower concentration is excitatory. Bozler's recent work, however, suggests that the inhibitory actions of adrenaline on the gastro-intestinal tract might eventually be correlated with some of its actions on skeletal muscle. His finding, for instance, that adrenaline has no effect on the intestinal pacemakers, or even on the propagation of excitation, but acts mainly by interfering with the *mechanism* of contraction (Bozler 1946) is not incompatible with the polarization and depression of propagation in skeletal muscle. His papers reveal, nevertheless, the extraordinary difficulties of working with tissues of such varied structural and functional types as smooth muscle organs and the grave problems of interpretation of electrical and mechanical phenomena in them.

Of recent years our knowledge of the action of acetylcholine on striated muscle and ganglion cells has increased very greatly, and we are perhaps within sight of

a reasonable explanation of its mode of action at these sites. But we have no clear picture of the way this substance acts when it is exerting its muscarine actions whether excitatory or inhibitory. A similar obscurity veils the effects of adrenaline on similar tissues. The time seems to me to be approaching when the elegant pharmacological techniques which Feldberg has described in this discussion must be supplemented by an attack on the processes of excitation and inhibition in smooth and cardiac muscle. I am only too fully aware of the difficulties confronting the electrophysiologist when dealing with cells of the size and complexity of those surrounding hollow organs, but the paucity of even the most elementary information holds out the promise of rich rewards to the experimenter with a sufficient temerity to begin such an investigation.

REMARKS ON CHEMICAL SPECIFICITY

By H. BLASCHKO

The phenomenon of chemical specificity is of importance not only in the study of the enzyme-substrate relationship, but also in the analysis of the specific action of the local hormones.

Substrate specificity of amine-forming enzymes

The idea that the biosynthesis of adrenaline requires a decarboxylation of an acid related to tyrosine is an old one (Friedmann 1906); it has recently gained experimental support by the demonstration that both the aromatic ring and the α-carbon atom of phenylalanine are incorporated in the adrenaline molecule (Gurin & Delluva 1947).

We have studied two amino-acid decarboxylases known to produce adrenaline-like amines. One is the mammalian enzyme L-DOPA decarboxylase (Holtz, Heise & Lüdtke 1938); it forms hydroxytyramine. The other is the L-tyrosine decarboxylase of *Streptococcus faecalis* R, which forms tyramine (Gale 1940; Epps 1944). This work started from the idea that the mammalian enzyme was the catalyst of one of the intermediate stages in the biosynthesis of adrenaline and sympathin (Blaschko 1939, 1942). It was found that the enzyme did not act on *N*-methyl-3:4-dihydroxy-phenylalanine (*N*-methyl-DOPA), a compound first described by Heard & Raper (1933). It has since been found that the inability to act on methylamino-acids is a general property of the bacterial as well as of the mammalian enzyme. The amino- and methylamino-acids tested are shown in table 7. In each case the amino-acid was not attacked.

TABLE 7

acid	corresponding amine	amine is formed	enzyme
tyrosine	tyramine	yes	bacterial
N-methyl tyrosine	*N*-methyl tyramine	no	
DOPA	hydroxytyramine	yes	mammalian
N-methyl DOPA	epinine	no	
dihydroxyphenylserine	noradrenaline	yes	mammalian and bacterial
adrenaline carboxylic acid	adrenaline	no	

The table shows that noradrenaline is formed by mammalian tissue extracts from dihydroxyphenylserine; the reaction is much slower than with the bacterial enzyme. For the bacterial enzyme it has been established that the amine formed was laevorotatory noradrenaline (Blaschko, Holton & Sloane Stanley 1948), and in recent experiments with Professor J. H. Burn and Dr H. Langemann the same has been found to be true for the mammalian extracts. We do not yet know if this reaction is important in the synthesis of noradrenaline (and adrenaline) in the animal body; other possible pathways remain to be investigated. Three substances which might be considered as immediate precursors of noradrenaline are shown in the diagram (figure 5).

$$HO\text{-}C_6H_3(OH)\text{-}CHOH\cdot CH(NH_2)\ COOH$$

$$HO\text{-}C_6H_4\text{-}CHOH\cdot CH_2\cdot NH_2 \qquad HO\text{-}C_6H_3(OH)\text{-}CH_2\cdot CH_2\cdot NH_2$$

$$HO\text{-}C_6H_3(OH)\text{-}CHOH\cdot CH_2\cdot NH_2$$
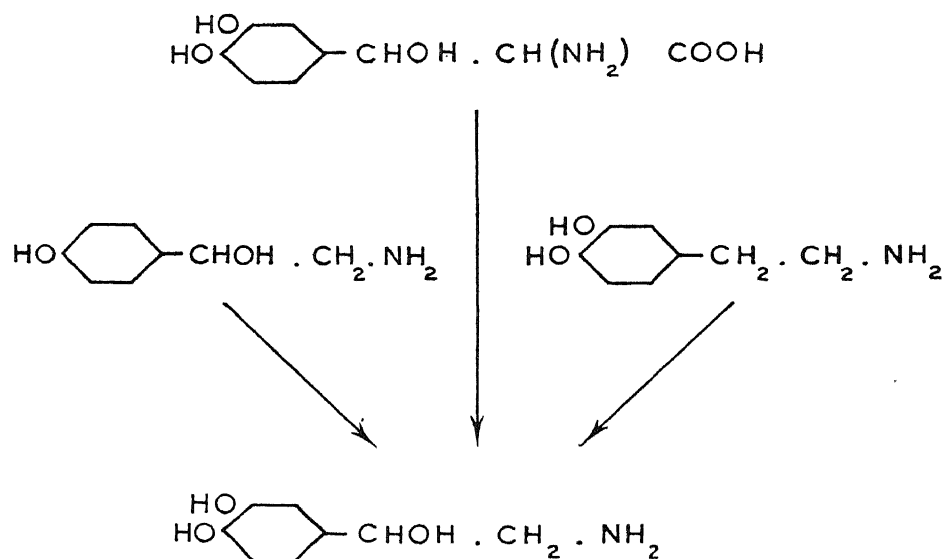
FIGURE 5.

We believe that the inability of the methylamino-acids to react with the two decarboxylases is explained by the formation of an intermediate compound between the coenzyme and the free amino group of the substrate (Blaschko 1950). The coenzyme is the same for both enzymes; it is pyridoxal phosphate. This is why enzymic activity in bacteria as well as in the rat depends upon the supply of pyridoxine in the culture medium (Umbreit, Bellamy & Gunsalus 1945) or in the diet (Blaschko, Carter, O'Brien & Sloane Stanley 1948).

There exist, however, differences in the substrate specificity of these closely related enzymes. This is illustrated by our observations on the decarboxylation of the isomers of tyrosine. It is known that DOPA decarboxylase does not act on tyrosine, but we have found that the enzyme will decarboxylate *m*-hydroxyphenylalanine and *o*-hydroxyphenylalanine (Blaschko, Holton & Sloane Stanley 1949; Blaschko 1949; Blaschko & Holton 1949). An acetone-dried preparation of *Streptococcus faecalis* R which contained an active tyrosine decarboxylase also acted on *m*-hydroxyphenylalanine, but it did not decarboxylate *o*-hydroxyphenylalanine at a significant rate (Sloane Stanley 1949).

The action of the bacterial enzyme on *m*-hydroxyphenylalanine presents an interesting difference from its action on tyrosine: with the acetone-dried preparation the rate of decarboxylation of *m*-hydroxyphenylalanine was about one-third of that of L-tyrosine. With the intact bacteria, however, the rate of decarboxylation of the meta-hydroxy isomer was almost negligible. This finding strongly suggests that the ability of the living bacterial cell to take up tyrosine is a highly selective property.

TABLE 8

	OH ⬡ CH$_2$ CHNH$_2$ COOH	⬡OH CH$_2$ CHNH$_2$ COOH	⬡OH CH$_2$ CHNH$_2$ COOH
substrate of			
mammalian enzyme	no	yes	yes
bacterial enzyme	yes	yes	no

These subtle differences in the substrate specificity must be due to differences in the two enzyme proteins. It seems tempting to assume that the substrate molecule reacts through the phenolic hydroxyl group with an atom in the enzyme protein, e.g. by forming a hydrogen bond (Blaschko 1950). This would explain why each of these two preparations will act on two isomers which have their hydroxyl groups in adjoining positions on the aromatic ring, and why they do not act on the third isomer.

The amino-acid decarboxylases show a strict specificity for L-amino-acids; they seem to be without affinity for the D-amino-acids. This is interesting in view of the results discussed. Of the four groups attached to the asymmetric α-carbon of the amino-acids, three are known to take part in the enzymic reaction:

$$\underset{\text{COOH}}{\overset{R}{\text{H—C—NH}_2}}$$
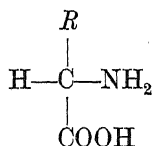
(*a*) the carboxyl group which loses its CO$_2$;

(*b*) the amino group which reacts with pyridoxal phosphate, and

(*c*) the group —*R* which reacts with the enzyme protein.

In order to account for the stereochemical specificity of these enzymes we must assume that these three groups take fixed positions in space relative to the enzyme. There can be only one configuration which fulfils this condition. At present we know least about the way in which the carboxyl group is orientated, but that it is essential in the decarboxylation reaction is obvious.

Can this discussion of stereochemical specificity be useful in understanding the observations on other enzymes? There exist two types of amino-acid oxidases, one

which acts on L-amino-acids, and another which is specific for D-amino-acids. These oxidases are much less affected than the decarboxylases by the nature of the group —R. This group can be modified within fairly wide limits; such alterations may affect the rate of oxidation, but not substrate specificity. For instance, all three DL-monohydroxyphenylalanines are substrates of the D-amino-acid oxidase of mammalian kidney as well as of the L-amino-acid oxidase of cobra venom. We have also examined the three corresponding DL-monochlorophenylalanines; none of these compounds is decarboxylated, but all three are substrates of the two amino-acid oxidases (Blaschko & Stiven 1950). Thus, for the reaction between the amino-acid molecule and the amino-acid oxidase the group —R seems to be less essential. How can we reconcile this fact with the high degree of stereochemical specificity of these enzymes? Clearly we can only do this by assuming that for the oxidase-substrate reaction the group —H is essential, and this is in fact well known to be true. The hydrogen atom in alpha position is believed to be one of the two which are transferred to the *iso*alloxazine ring of flavine-adenine-dinucleotide in the dehydrogenation reaction (see Krebs 1948); its replacement by a methyl group results in a compound which is not a substrate of the mammalian D-amino-acid oxidase (Keilin & Hartree 1936).

The stereochemical specificity of laevo-adrenaline

Can similar concepts be useful in our understanding of the specific action of local hormones? The adrenaline molecule has an asymmetric carbon atom; its laevorotatory isomer which occurs in nature, is highly active. We shall therefore assume that three of the groups attached to the asymmetric carbon atom are essential in the reaction with the excitable tissue. We assume that these three groups are:

(a) the catechol group,

(b) the group —$CH_2.NH.CH_3$, and

(c) the hydroxyl group.

These three groups are assumed to be arranged in a fixed spatial relationship relative to the tissue receptors. We might say: the tissue receptors for adrenaline have three receptacles or anchorages, one for each of these groups.

The stereochemical specificity of the action of adrenaline differs from that of the enzyme discussed above; the specificity is not absolute: dextro-adrenaline has about $\frac{1}{12}$ to $\frac{1}{15}$ of the action of laevo-adrenaline on the arterial blood pressure. How can this be explained? It is obvious that dextro-adrenaline cannot attach itself in the same way to the receptors as laevo-adrenaline. We must assume that one of the three receptacles is not engaged: the dextro-adrenaline molecule is attached by the basic group and by the catechol group; the hydroxyl group is not engaged. The receptacle for the hydroxyl group is faced by the hydrogen atom attached to the asymmetric carbon atom. The dextro-adrenaline molecule is therefore held in two points; this attachment is less firm; the dextro-adrenaline molecule has a greater degree of freedom on the receptor and it is therefore less active.

Can this interpretation of the mode of attachment of the adrenaline molecule be supported by observation? It has been pointed out by Barger & Dale (1910)

that the presence of the catechol group and the basic group are essential for adrenaline-like action. They found that the compound methylaminoethyl catechol (also known as epinine) was adrenaline-like in its action. On the arterial blood pressure it has about $\frac{1}{12}$ of the action of laevo-adrenaline. This is in agreement with what we would expect: on the hypothesis outlined, epinine turns towards the receptor a face that looks like that of dextro-adrenaline. In other words, the catechol group and the basic group mainly determine the quality of the response; the hydroxyl group, provided it is present in the right position, mainly determines the intensity of the attachment.
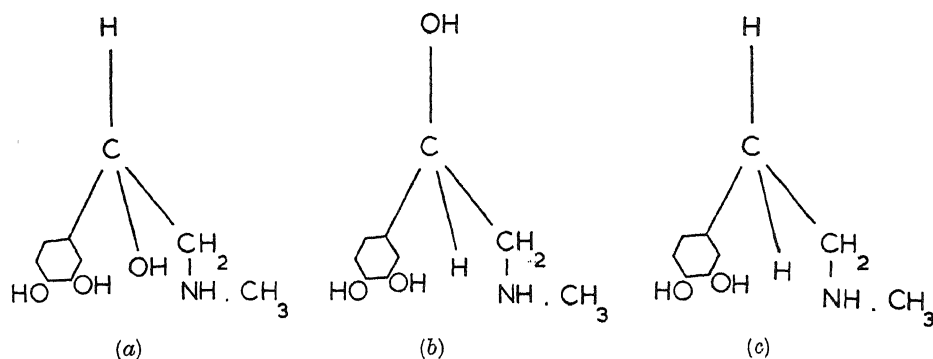
FIGURE 6. The attachment of (a) laevo-adrenaline, (b) dextro-adrenaline, and (c) epinine to the adrenergic receptors. The three groups which are essential for attachment are those drawn below the carbon atom.

This discussion is not the place to enquire if this concept of stereochemical specificity is more widely applicable, but it seems likely that it will be found useful wherever differences are observed in substances in which stereoisomerism is due to the presence of an asymmetric carbon atom.

CHOLINESTERASE ACTIVITY AND THE INNERVATION OF PERIPHERAL TISSUES

BY R. H. S. THOMPSON

Although the nature of the cholinesterases in brain, red blood cells and plasma has been studied extensively, little comparative information is available on the distribution and characteristics of the types of cholinesterase present in other tissues, and I would like to describe very briefly some of the results which Miss M. G. Ord and I have been obtaining in the course of a survey of different mammalian tissues (Ord & Thompson 1950*a*).

Owing to the species differences in blood esterases we have limited our work to a study of rat and human tissues. In order to obtain some initial information about the distribution of the so-called 'true' and 'pseudo-' cholinesterases we began by comparing the relative rates of hydrolysis of acetylcholine (ACh) and of acetyl-β-methylcholine (MCh) and benzoylcholine (BCh); that is to say, we used

the physiological substrate and also the selective substrates for the true and pseudo-cholinesterases introduced by Mendel, Mundell & Rudney (1943). Apart from a few scattered observations on a small number of tissues from different animal species no attempt to compare systematically the distribution of these esterases in different tissues of animals of the same species was made until Sawyer & Everett (1947) showed that the liver, the uterus and a number of glandular tissues of the rat were able to hydrolyze benzoylcholine more rapidly than acetyl-β-methylcholine.

In our experiments with rats, after killing the animals by decapitation and exsanguination, the carcasses were perfused with 0·9% NaCl until the effluent fluid was no longer visibly blood-stained, after which the tissues under study were removed, homogenized in 0·025M-NaHCO$_3$ and the cholinesterase activity in the presence of the appropriate substrate measured manometrically. In the experiments with skin the tissue was sliced.

The results that we obtained are summarized in table 9. The activity of the homogenates of the different tissues is expressed as μl. CO$_2$/g. tissue (wet wt.)/hr., while the last two columns show the activities in the presence of the two selective substrates expressed as a percentage of the activity in the presence of acetyl choline.

TABLE 9. RELATIVE CHOLINESTERASE ACTIVITIES OF RAT TISSUES TOWARDS ACh, BCh AND MCh

		activity (μl. CO$_2$/g./hr.)			activity as percentage of ACh activity	
group	tissue	ACh (0·015M)	BCh (0·015M)	MCh (0·03M)	BCh	MCh
A	brain	5,535	346	4,245	6	75
	quadriceps femoris	231	16	158	7	69
	diaphragm	628	0	342	0	54
	suprarenal gland	2,540	47	994	2	39
B	stomach	1,535	367	309	24	20
	liver	912	212	186	23	20
	lung	848	215	175	25	21
	salivary gland	1,589	489	523	31	33
C	heart ventricle	3,600	1,050	308	29	9
	heart auricle	10,780	3,400	930	32	9
	intestinal muscle	3,670	1,517	615	41	17
	intestinal mucosa	6,460	3,255	468	50	7
	Harderian gland	6,210	3,250	401	52	6
	skin	3,030	762	294	25	9

It will be seen that, as might be expected, there are wide variations in the rates of hydrolysis of acetylcholine by different tissues. But the relative rates of hydrolysis of BCh and of MCh also differ strikingly, and the tissues of this small series which we have studied appear to fall into three groups:

Group A. Tissues which hydrolyze MCh very much more rapidly than BCh, the hydrolysis rate of BCh amounting in fact to only 0 to 6 % of the rate at which acetylcholine is hydrolyzed. To use Mendel's interpretation this group therefore contains almost exclusively the true cholinesterase.

Group B. Tissues which contain both the true and pseudo-cholinesterases in such amounts that BCh and MCh are hydrolyzed at roughly equal rates.

Group C. Tissues which contain predominantly a pseudo-cholinesterase.

Further evidence in support of the view that two distinct enzymes are concerned with the hydrolysis of BCh and MCh by the tissues of groups B and C was obtained by carrying out summation experiments in the presence of both substrates; we found that there was an additive effect, the CO_2 output in the presence of both substrates amounting to 81 to 102 % of the sum of the CO_2 outputs in the presence of each substrate separately.

It was also shown that the enzyme hydrolyzing BCh in these tissues is of the same type as the pseudo-cholinesterase described in the serum of various animal species, and is not related to the more specific 'benzoylcholine esterase' described in guinea-pig liver. For example, the enzyme in salivary gland, lung or heart ventricle is highly sensitive to inhibition by di-*iso*propyl fluorophosphonate (DFP), concentrations of the order of 1×10^{-8}M causing 50 % inhibition, while concentrations of about 5×10^{-8}M cause virtually complete inhibition. The enzyme is also sensitive to inhibition by eserine and is active against butyrylcholine, which has been shown to be another selective substrate for the pseudo-cholinesterase (Easson & Stedman 1936–7; Adams & Whittaker 1949).

One other point of interest in connexion with the innervation of these tissues is that by contrast with brain, skeletal muscle and suprarenal gland, the cells of those tissues which exhibit an activity towards BCh either equalling or exceeding that towards MCh all receive only post-ganglionic innervation. Also, the very high acetylcholine-hydrolyzing activities shown by heart auricle, intestinal mucosa and Harderian gland are of interest, particularly the case of the intestinal mucosa, in view of the finding by Feldberg & Lin (1949 *b*) of a relatively high value of choline acetylase in the glandular mucosa of the duodenum and ileum, a layer usually regarded as devoid of nerve cells. We cannot, of course, conclude from our present observations anything as to the physiological significance of the pseudo-cholinesterase in these tissues; we have, however, shown that the high acetylcholine-hydrolyzing activity of, for example, heart auricle or intestinal mucosa is due mainly to the presence of a pseudo-cholinesterase which according to some workers is unconnected with nervous activity. It is tempting therefore to speculate on the possibility that the pseudo-cholinesterase in these sites may be concerned with destroying acetylcholine liberated as part of some non-nervous, local controlling function.

To turn to human tissues, these results are rather less complete than those with rats, but we have been able to show that in man there is also an active pseudo-cholinesterase in certain peripheral muscular organs possessing only post-ganglionic innervation, namely, uterus, ureter, jejunal muscle and jejunal mucosa. Again, in man, the jejunal mucosa is about twice as active as the jejunal muscle.

The human tissue which we have studied most extensively is the placenta. Torda (1942) had previously shown that perfused blood-free placental tissue hydrolyzes acetylcholine, but she did not report on the specificity of the enzyme in this tissue.

We have compared the rates of hydrolysis of acetylcholine, acetyl-β-methylcholine and benzoylcholine by both fresh and perfused human placenta (Ord & Thompson 1950b). From table 10 it will be seen that while fresh placenta with its contained blood hydrolyzes both BCh and MCh, the perfused blood-free placenta shows only a very low order of activity in the presence of benzoylcholine, although it still hydrolyzes acetyl-β-methylcholine at an appreciable rate.

TABLE 10. HYDROLYSIS RATES OF ACh, MCh AND BCh BY FRESH AND
PERFUSED HUMAN PLACENTA

		activity (μl. CO_2/g./hr.)		
		ACh	MCh	BCh
fresh		420	169	50
		412	198	68
		1110	370	239
		742	284	173
	mean	671	255	132
perfused		155	66	6
		175	121	11
		274	162	16
		220	121	24
		186	84	0
	mean	202	111	11

Using Mendel's terminology again, therefore, it would seem that perfused placenta contains predominantly a true cholinesterase. Further evidence supporting the view that the placental enzyme is of this type has been obtained by studying its sensitivity to inhibition by DFP; we have found that the concentration required to cause 50 % inhibition (I_{50} concentration) is $1 \cdot 0 \times 10^{-7}$M for the placental enzyme, which agrees well with the I_{50} value of $1 \cdot 3 \times 10^{-7}$M for the enzyme in human brain (Adams & Thompson 1948).

Placenta, like the red blood cells, provides therefore another example of a nerve-free tissue which contains almost exclusively the true cholinesterase. Bullock, Grundfest, Nachmansohn & Rothenberg (1947) have stated that of a great variety of tissues tested, nerve and muscle were the only ones in which a 'specific' cholinesterase could be tested. Our finding of a true cholinesterase in placenta, apparently similar to the specific enzyme in brain, suggests therefore that caution should be exercised in associating this one type of cholinesterase specifically or exclusively with nervous activity, as has been done by some workers. It seems possible that in the placenta this true cholinesterase may be playing a part in connexion with some local controlling action of acetylcholine, and it is of interest that, if this is so, the destruction of acetylcholine in this nerve-free tissue should be carried out by an enzyme apparently similar to the one in nervous tissue.

By H. J. C. Swan

Dr H. J. C. Swan described experiments in which he studied the effect of intra-venous infusion of adrenaline on the blood flow in the hand of man. During the infusion he observed a fall in the rate of blood flow. When the infusion stopped at the end of 3 min. he observed an increase in the rate of blood flow. This increase of flow was not observed if the hand was one in which the sympathetic fibres had degenerated after sympathectomy. Moreover, this increase was not observed if the infusion was made by the intra-arterial route. The same observations were made with adrenaline and at times with noradrenaline.

By N. Ambache

A neurological factor, which is usually overlooked, should be mentioned, as it may provide an explanation of certain inhibitory phenomena, referred to earlier in the discussion, which are observed both in physiological and in pharmacological experiments on mammalian intestines. For various reasons the existence of a distinct set of inhibitory ganglion cells in the gut seems highly probable. This may have been in Langley's mind when he published (Langley 1922) a diagram of the connexions of the enteric nerve cells, which presents several interesting features. For instance, among the cells labelled by him 'local cells', which are all the ultimate cells directly innervating the smooth muscle, he has figured alternate postganglionic neurones as motor and inhibitory.

A moment's reflection will show that at least this part of the diagram is probably correct, for Bayliss & Starling (1899) showed that the peristaltic reflex comprises two distinct components. There is a motor component above the bolus, and an inhibitory component, consisting of active relaxation of the gut, below it. Viewed in this way, the organization of the peristaltic reflex centres provides yet another example of reciprocal innervation. Certain histological observations by Kuntz (1922) are consistent with this concept of the duality of the enteric plexus. Among the types of cells found by him in enteric ganglia, two were distinct both in their staining properties and in their mutual relationship. In figures 4 to 7 of his paper there are four examples of what is stated (on p. 198 of that paper) to be a darker staining nerve cell, sending a process which ends in synaptic formations about a more lightly stained cell. It is quite possible that this arrangement typifies the reciprocal connexions in these nerve centres, and that one of these two neurons is inhibitory.

There is also reason to suppose that the inhibitory nerve cells in the enteric plexus are accessible to vagal impulses in certain species (cats and dogs). For it is known that, under certain conditions, e.g. after atropine, vagal stimulation can produce inhibition in the gut (Bayliss & Starling 1899, 1901; McSwiney & Robson 1929). It is, presumably, in order to explain these inhibitory effects of vagal stimulation, that Langley marked some of the vagal preganglionic fibres in his diagram with a minus sign. However, there is no evidence at present to support the concept of an inhibitory function of preganglionic fibres in the autonomic system. For, although inhibition of the second of two successive responses does

occur in some ganglionic synapses (Eccles 1935; Whitteridge 1937; Lloyd 1939), this could not account for the reversal of the effect on the muscle. For instance, although a type of inhibition can be demonstrated in the ciliary ganglion (Whitteridge 1937), stimulation of the oculomotor nerve does not, in the absence of stimulus spread to sympathetic fibres, dilate the pupil in mammals. For these reasons it is felt that Langley's diagram were better amended to show an excitatory connexion (existing in cats and dogs) between the vagal preganglionic fibres and the inhibitory ganglion cells in the gut; this connexion may be via the reciprocal arrangement described above. Thus, vagal impulses may be presumed to impinge simultaneously upon both motor and inhibitory neurones in the gut. Normally the motor effect is predominant, although there are signs of underlying inhibition (Bayliss & Starling 1899, 1901); after atropine, the cholinergic motor effect is reduced, and the simultaneous stimulation of the inhibitory nerve cells is revealed.

Since these inhibitory ganglion cells are participant in synapses, it may be inferred that they are responsive to nicotine and to other ganglion-cell stimulating substances. And, indeed, signs of an underlying inhibitory process are sometimes seen in normal intestinal preparations, immediately after a small stimulant dose of nicotine is washed out of the organ bath. Further evidence of this has been obtained recently in intestinal preparations poisoned with botulinum toxin. It is now known that this toxin paralyses those nerve endings which are cholinergic, sparing the non-cholinergic ones; it is therefore convenient for differential 'denervation' of an effector tissue (in this case for the selective paralysis of the motor endings in the gut). In intestinal preparations from rabbits injected locally with a lethal dose of toxin 6 to 7 hr. previously, pure inhibitory responses were elicited by small doses of nicotine, which invariably contract a normal preparation. The final concentration of nicotine in these experiments was very low, varying from 0·2 to 2×10^{-6}, and it is unlikely that even much larger amounts of nicotine would have a depressing effect on muscle fibres, in contradistinction to a suggestion made earlier in this discussion. For example, in experiments on the sheep's uterus, Ambache & Hammond (1946 unpublished) found that the uterine muscle could tolerate 25 to 100 times that amount of nicotine without showing signs of depression. The same deduction can be made from the results obtained by Thomas & Kuntz (1926), who described the effect upon the intestinal response to vagal stimulation, of increasing doses of nicotine. The response was at first reduced by 0·3 mg./kg. owing to, as they suggest, a stimulation of sympathetic ganglion-cells. Later, there was recovery in the presence of 50 mg./kg.; potentiation by 736 mg./kg.; and, finally, extinction only at a level of 2000 to 3000 mg./kg., a truly enormous dose.

The presence of inhibitory elements in the gut itself opens up the possibility of yet another type of 'indirect' pharmacological action on smooth muscle. It may perhaps enable us to understand certain pharmacological oddities observed by Bernheim (1931). This author found that the addition of (1/1 500 000) pilocarpine, which is a stimulator of ganglion cells (see Ambache 1949), to a preparation of guinea-pig's intestine (jejunum or duodenum; not ileum) at the height of a histamine-contraction, produced relaxation; further addition of histamine, after the onset of this relaxation, was without effect, unless atropine (which is known to

block the action of pilocarpine on ganglion cells (Ambache 1949) was first added. When given alone, pilocarpine first contracted the gut, and then relaxed it, whilst still in the bath. Here, then, is an instance of an inhibitory effect of pilocarpine; yet there has been no suggestion so far that this substance is a muscle-fibre depressant.

In conclusion, the possibility should be taken into account that some types of pharmacological inhibitions in the gut could arise merely through the activation of inhibitory ganglion cells.

CONCLUSION

In concluding the discussion, the opener, in reply to the question of Professor Bacq whether choline would cause resumption of the beat of the arrested auricles, said that concerning the auricles which stopped spontaneously he did not know, but concerning the auricles stopped by paludrine, choline would cause the beat to start again if given in 1000 times the amount of acetylcholine. In reply to Professor Brown, who asked whether there was any evidence for reversal of the action of the vagus, the opener said that there was. Dale, Laidlaw and Symons had shown that in a cat under paraldehyde anaesthesia, after the injection of extract of curare, or of nicotine, or of tropine, the stimulation of the vagus nerve caused acceleration of the heart rate.

REFERENCES

Abdon, N.-O. & Hammarskjöld, S. O. 1944 *Acta physiol. Scand.* **8**, 75.
Adams, D. H. & Thompson, R. H. S. 1948 *Biochem. J.* **42**, 170.
Adams, D. H. & Whittaker, V. P. 1949 *Biochim. Biophys. Acta*, **3**, 368.
Ahlquist, R. P. 1948 *Amer. J. Physiol.* **153**, 586.
Ambache, N. 1946 *J. Physiol.* **104**, 266.
Ambache, N. 1949 *J. Physiol.* **110**, 145, 164.
Bacq, Z. M. 1933 *Arch. int. Physiol.* **36**, 167.
Bacq, Z. M. 1947 *Biol. Rev.* **22**, 73.
Bacq, Z. M. 1949 *J. Pharmacol.* **95** (Part II), 1.
Bacq, Z. M. & Goffart, M. 1939 *Arch. int. Physiol.* **49**, 179.
Bacq, Z. M. & Monnier, A. M. 1935 *Arch. int. Physiol.* **40**, 467.
Bacq, Z. M. & Weekers, R. 1949a *Bull. Acad. Roy. Med. Belg.* **14**, 183.
Bacq, Z. M. & Weekers, R. 1949b *Bull. Acad. Roy. Med. Belg.* **14**, 299, 484.
Barger, G. & Dale, H. H. 1910 *J. Physiol.* **41**, 19.
Bartlett, G. R. & Barron, E. S. G. 1947 *J. Biol. Chem.* **170**, 67.
Bayliss, W. M. & Starling, E. H. 1899 *J. Physiol.* **24**, 99.
Bayliss, W. M. & Starling, E. H. 1901 *J. Physiol.* **26**, 125.
Bernheim, F. 1931 *J. Pharmacol.* **43**, 509.
Beznak, A. B. L. 1936 *Quart. J. Exp. Physiol.* **26**, 253.
Blaschko, H. 1939 *J. Physiol.* **90**, 50P.
Blaschko, H. 1942 *J. Physiol.* **101**, 337.
Blaschko, H. 1949 *Biochem. J.* **44**, 268.
Blaschko, H. 1950 *Biochim. Biophys. Acta* (Meyerhof Anniversary Issue), **4**, 130.
Blaschko, H., Carter, C. W., O'Brien, J. R. P. & Sloane Stanley, G. H. 1948 *J. Physiol.* **107**, 18P.
Blaschko, H. & Holton, P. 1949 *J. Physiol.* **110**, 482.
Blaschko, H., Holton, P. & Sloane Stanley, G. H. 1948 *Brit. J. Pharmacol.* **3**, 315.
Blaschko, H., Holton, P. & Sloane Stanley, G. H. 1949 *J. Physiol.* **108**, 427.
Blaschko, H. & Stiven, J. 1950 *Biochem. J.* **46**, 88.
Bonnet, V. 1938 *Arch. int. Physiol.* **47**, 397.
Bozler, E. 1941 *Cold Spr. Harb. Symp. Quant. Biol.* **3**, 95.

Bozler, E. 1946 *Amer. J. Physiol.* **146**, 496.

Brink, F. Jr., Bronk, D. W. & Larabee, M. G. 1946 *Ann. N.Y. Acad. Sci.* **47**, art. 4, 457.

Brouha, L. & Bacq, Z. M. 1936 *Arch. int. Physiol.* **43**, 492.

Brown, G. L. 1937 *J. Physiol.* **89**, 438.

Brown, G. L., Bülbring, E. & Burns, B. D. 1948 *J. Physiol.* **107**, 115.

Brown, G. L., Burns, B. D. & Feldberg, W. 1948 *J. Physiol.* **107**, 346.

Brown, G. L. & Feldberg, W. 1936 *J. Physiol.* **88**, 265.

Brown, G. L., Goffart, M. & Vianna Dias, M. 1950 *J. Physiol.* **111**, 184.

Brown, G. L. & Harvey, A. M. 1938 *J. Physiol.* **94**, 101.

v. Brücke, E. T. & Oinuma, S. 1910 *Pflüg. Arch. ges. Physiol.* **136**, 502.

Buchthal, F. & Lindhard 1942 *Acta physiol. Scand.* **4**, 2.

Buffa, P. & Peters, R. A. 1949 *J. Physiol.* **110**, 488.

Bülbring, E. 1944 *J. Physiol.* **103**, 55.

Bülbring, E. & Burn, J. H. 1949*a* *J. Physiol.* **108**, 508.

Bülbring, E. & Burn, J. H. 1949*b* *J. Physiol.* **108**, 6*P*.

Bülbring, E., Lourie, E. M. & Pardoe, A. U. 1949 *Brit. J. Pharmacol.* **4**, 290.

Bullock, T. H., Grundfest, H., Nachmansohn, D. & Rothenberg, M. A. 1947 *J. Neurophysiol.* **10**, 11.

Burn, J. H. & Vane, J. R. 1949 *J. Physiol.* **108**, 104.

Burns, B. D., Paton, W. D. M. & Vianna Dias, M. 1950 *Proc. Colloque Sur Facteurs Ioniques* (Paris) (in the Press).

Cannon, W. R. & Rosenblueth, A. 1937 *Autonomic Neuro-effector systems.* New York: Macmillan Co.

Chang, H. C. & Gaddum, J. H. 1933 *J. Physiol.* **79**, 255.

Chèvremont, M. 1945 *J. Morphol.* **76**, 139.

Chèvremont, M. 1948 *Biol. Rev.* **23**, 267.

Clark, A. J. 1937 General Pharmacology. *Heffters Handbuch. exp. Pharm. akol.* Ergänz. 4.

Collins, D. A. 1948 *J. Pharmacol.* **94**, 244.

Comline, R. S. 1946 *J. Physiol.* **105**, 6*P*.

Coppée, G. & Bacq, Z. M. 1938 *Arch. int. Physiol.* **47**, 312.

De Robertis, E. & Schmitt, F. O. 1949 *J. Exp. Med.* **90**, 283.

Derouaux, G. 1941 *Arch. int. Pharmacodyn.* **66**, 202.

Dews, P. B. & Graham, J. D. P. 1946 *Brit. J. Pharmacol.* **1**, 287.

Donomae, E. & Feldberg, W. 1934 *Pflüg. Arch. ges. Physiol.* **147**, 171.

Draper, M. H. & Weidmann, S. 1950 Personal communication.

Easson, L. H. & Stedman, E. 1936–7 *Proc. Roy. Soc.* B, **121**, 142.

Eccles, J. C. 1935 *J. Physiol.* **85**, 207.

Eccles, J. C. & Magladery, J. W. 1937*a* *J. Physiol.* **90**, 31.

Eccles, J. C. & Magladery, J. W. 1937*b* *J. Physiol.* **90**, 68.

Einthoven, W. & Rademaker, A. C. A. 1917 *Pflüg. Arch. ges. Physiol.* **166**, 109.

Elliott, T. R. 1904 *J. Physiol.* **31**, xx.

Emmelin, N. & Feldberg, W. 1947 *J. Physiol.* **106**, 482.

Epps, H. M. R. 1944 *Biochem. J.* **38**, 242.

Euler, U. S. v. 1946 *Acta Physiol. Scand.* **11**, 168; **12**, 73.

Euler, U. S. v. 1948*a* *Acta Physiol. Scand.* **16**, 63.

Euler, U. S. v. & Astrom, A. 1948*b* *Acta physiol. Scand.* **16**, 97.

Feldberg, W., Harris, G. & Lin, R. C. Y. 1950 (Unpublished experiments.)

Feldberg, W. & Lin, R. C. Y. 1949*a* *J. Physiol.* **109**, 475.

Feldberg, W. & Lin, R. C. Y. 1949*b* *J. Physiol.* **109**, 32*P*.

Feldberg, W. & Lin, R. C. Y. 1950 *J. Physiol.* **111**, 96.

Feldberg, W. & Mann, T. 1946 *J. Physiol.* **104**, 411.

Feldberg, W. & Rosenfeld, P. 1933 *Pflüg. Arch. ges. Physiol.* **232**, 212.

Feldberg, W. & Schilf, E. 1930 *Histamin.* Berlin: Springer.

Fischer, E. 1944 *Physiol. Rev.* **24**, 467.

Fischer, P. & Lecomte, J. 1950 *Arch. int. Physiol.* (in the Press).

Fredericq, H. 1924 *Arch. int. Physiol.* **23**, 168.

Fredericq, H. 1925*a* *Arch. int. Physiol.* **24**, 295.

Fredericq, H. 1925*b* *C.R. Soc. Biol., Paris,* **92**, 208.

Fredericq, H. 1925*c* *C.R. Soc. Biol., Paris,* **92**, 739.

Fredericq, H. 1926 *C.R. Soc. Biol., Paris*, **95**, 247.

Fredericq, H. 1931 *Bull. Acad. Roy. Med. Belg.* **11**, 272.

Fredericq, H. 1938 *Arch. int. Physiol.* **46**, 195.

Friedenwald, J. S. & Buschke, W. 1944 *Amer. J. Physiol.* **141**, 689.

Friedmann, E. 1906 *Beitr. chem. Physiol. Path.* **8**, 95.

Gaddum, J. H. 1936 *Gefässerweiternde Stoffe der Gewebe*. Leipzig: Thieme.

Gaddum, J. H. & Goodwin, L. G. 1947 *J. Physiol.* **105**, 357.

Gaddum, J. H. & Kwiatkowski, H. 1938 *J. Physiol.* **94**, 87.

Gaddum, J. H. & Kwiatkowski, H. 1939 *J. Physiol.* **96**, 385.

Gale, E. F. 1940 *Biochem. J.* **34**, 392.

Gaskell, J. F. 1887 *J. Physiol.* **8**, 404.

Goffart, M. & Perry, W. L. M. 1950 *J. Physiol.* (in the Press).

Govaerts, J. 1939 *Arch. int. Physiol.* **49**, 426.

Green, D. E. & Richter, D. 1937 *Biochem. J.* **31**, 596.

Greer, C. M., Pinkston, J. O., Baxter, J. H. & Brannon, E. S. 1938 *J. Pharmacol.* **62**, 189.

Gurin, S. & Delluva, A. M. 1947 *J. Biol. Chem.* **170**, 545.

Hambleton, B. F. 1914 *Amer. J. Physiol.* **34**, 466.

Heard, R. D. H. & Raper, H. S. 1933 *Biochem. J.* **27**, 36.

Heymans, C. 1949*a* *Bull. Acad. Roy. Med. Bel.* **14**, 379.

Heymans, C. 1949*b* *Bull. Acad. Roy. Med. Belg.* **14**, 383.

Holton, P. 1949 *Nature*, **163**, 217.

Holton, P. 1950 (Quoted from Feldberg & Lin 1950.)

Holtz, P., Heise, R. & Lüdtke, K. 1938 *Arch. exp. Path. Pharmak.* **191**, 87.

Irwin, D. A. 1931 *Amer. J. Anat.* **49**, 141.

Jang, C. S. 1940 *J. Pharmacol.* **70**, 347; **71**, 87.

Keilin, D. & Hartree, E. F. 1936 *Proc. Roy. Soc. B*, **119**, 114.

King, C. E. & Arnold, L. 1922 *Amer. J. Physiol.* **59**, 97.

Kolm, R. & Pick, E. P. 1921 *Pflüg. Arch. ges. Physiol.* **190**, 108.

Konzett, H. 1940 *Arch. exp. Path. Pharmak.* **197**, 27.

Krebs, H. A. 1948 *Biochem. Soc. Symp.* **1**, 2.

Kuffler, F. 1943 *J. Neurophysiol.* **6**, 99.

Kuntz, A. 1922 *Anat. Rec.* **24**, 193.

Langley, J. N. 1922 *J. Physiol.* **56**, 39*P*.

Lardy, H. A. 1949 *Respiratory enzymes*. Minneapolis: Burgess Publishing Company.

Le Heux, J. W. 1918–19 *Pflüg. Arch. ges. Physiol.* **173**, 8.

Lettre, H. 1943 *Chem. Z.* **67**, 52.

Liébecq, C. & Peters, R. A. 1949 *Biochim. Biophys. Acta*, **3**, 215.

Lloyd, D. P. C. 1939 *J. Physiol.* **95**, 464.

Loewi, O. 1921 *Pflüg. Arch. ges. Physiol.* **189**, 239.

Loewi, O. 1936 *Pflüg. Arch. ges. Physiol.* **237**, 504.

McCrea, E. D., McSwiney, B. A. & Stopford, J. S. B. 1925 *Quart. J. Exp. Physiol.* **18**, 301.

McDowall, R. J. S. 1946 *J. Physiol.* **104**, 392.

MacIntosh, F. C. 1938 *Arch. int. Physiol.* **47**, 321.

McSwiney, B. A. & Robson, J. M. 1929 *J. Physiol.* **68**, 124.

McSwiney, B. A. & Wadge, W. J. 1928 *J. Physiol.* **65**, 350.

Maddock, W. O., Rankin, V. M. & Youmans, W. B. 1948 *Proc. Soc. Exp. Biol., N.Y.*, **67**, 151.

Magnus, R. 1920 *Naturwissenschaften*, **8**, 383.

Martius, C. 1949 *Liebigs Ann.* **561**, 227.

Mendel, B., Mundell, D. B. & Rudney, H. 1943 *Biochem. J.* **37**, 473.

Monnier, A. M. 1936 *Cold Spr. Harb. Symp. Quant. Biol.* **4**, 111.

Monnier, A. M. & Dubuisson, M. 1934*a* *Arch. int. Physiol.* **38**, 180.

Monnier, A. M. & Dubuisson, M. 1934*b* *Arch. int. Physiol.* **38**, 207.

Nasmyth, P. A. 1950 *J. Physiol.* **110**, 294.

Ord, M. G. & Thompson, R. H. S. 1950*a* *Biochem. J.* **46**, 346.

Ord, M. G. & Thompson, R. H. S. 1950*b* *Nature* (in the Press).

Paton, W. D. M. & Zaimis, E. J. 1949 *Brit. J. Pharmacol.* **4**, 381.

Peart, W. S. 1949 *J. Physiol.* **108**, 491.

Prosser, C. L. 1942 *Biol. Bull. Woods Hole*, **83**, 145.

Richter, D. & Tingey, A. H. 1939 *J. Physiol.* **97**, 265.

Sachs, A. 1937 *Cardiologia*, 1, 74.
Saunders, B. C. 1947 *Nature*, 160, 179.
Sawyer, C. H. & Everett, J. W. 1947 *Amer. J. Physiol.* 148, 675.
Schmiterlöw, C. G. 1948 *Acta physiol. Scand.* 16, Suppl. 56.
Sloane Stanley, G. H. 1949 *Biochem. J.* 44, 373.
Spadolini, I. & Domini, G. 1940 *Arch. Fisiol.* 40, 147.
Thomas, J. A. 1949 *Rev. Hématol.* 4, 639.
Thomas, J. E. & Kuntz, A. 1926 *Amer. J. Physiol.* 76, 598.
Thompson, R. H. S. 1950 Unpublished data.
Torda, C. 1942 *Proc. Soc. Exp. Biol., N.Y.*, 51, 398.
Tullar, B. F. 1949 *Science*, 109, 536.
Umbreit, W. W., Bellamy, W. D. & Gunsalus, I. D. 1945 *Arch. Biochem.* 7, 185.
Verzar, F. & v. Kokas, E. 1927 *Pflüg. Arch. ges. Physiol.* 217, 397.
Weiland, W. 1912 *Pflüg. Arch. ges. Physiol.* 147, 171.
West, G. B. 1947 *J. Physiol.* 106, 418.
Whitteridge, D. 1937 *J. Physiol.* 89, 99.
Wright, R. D., Jennings, M. A., Florey, H. W. & Lium, R. 1940 *Quart. J. Exp. Physiol.* 30, 72.
Zeller, E. A., Schär, B. & Staehlin, S. 1939 *Helv. chim. Acta*, 22, 837.

The development of the active state of muscle during the latent period

By A. V. Hill, F.R.S., *Biophysics Research Unit, University College, London*

The mechanical response of muscle, as ordinarily recorded, depends on the shortening of its contractile elements, which is a slow process. An applied stretch obviates the necessity of shortening and allows the change of state resulting from stimulation to be examined directly. During the latent period a considerable mechanical change occurs, the muscle rapidly becoming less extensible. This change begins about half-way through the latent period; by its end a very substantial resistance to stretch is encountered.

The effect described is much greater than the 'latency relaxation', though beginning at about the same moment. The 'latency relaxation' is probably due to a slight increase of length accompanying the major physical or molecular change made evident by stretching. The early increase of transparency may be a sign of the same process.

The rapid decrease of extensibility has been followed beyond the latent period into the earlier stages of contraction.

The decrease of extensibility begins at about the same moment as the heat of activation; the latter is regarded as a product of the chemical process by which the change of mechanical state is effected.

In experiments described in a recent paper (Hill 1949*e*) the mechanical condition of an active muscle was examined by applying a rapid stretch at various moments after a single maximal shock. It was concluded that at the end of the latent period there is an abrupt change of state, the contractile component of the muscle suddenly becoming capable of bearing a considerable load. The intensity of the active state, defined as the maximum force which the muscle can bear without lengthening, was greatest at the start, was maintained for a time and then declined as relaxation set in.

No change of this kind can really be abrupt, it must occur in a finite interval, possibly too short for accurate investigation. The muscles employed, the sartorii of

frog and toad, were too rapid to allow a closer analysis in time of the onset of the active state. Such an analysis would require a very quick stretch which, with the instruments available, would cause vibration and obscure the records. It was desirable, however, for three reasons, to explore the mechanical condition of a muscle during the latent period itself:

(1) Sandow in a series of papers (e.g. 1944, 1947) has shown that, at an interval after a shock considerably less than the latent period as ordinarily observed, a muscle initially under tension shows a small initial relaxation ('latency relaxation') which is rapidly reversed by the ensuing contraction; this has been confirmed at 0° C by Abbott & Ritchie (1948, 1950), who found in frogs' muscles that the tension drop begins at about 7·5 msec., being reversed at about 18 msec., the usual latent period.

(2) D. K. Hill (1949) has shown that during the twitch of a frog's skeletal muscle, starting during the latent period at the same moment as the 'latency relaxation', there is a change in optical transparency. An early rapid phase (increased transparency), occurring during the 'latency relaxation', has certain characteristics which suggest that its cause is essentially different from that of the later phase (also increased transparency) which accompanies contraction proper and is complete (von Muralt 1934) only after recovery has restored the muscle to its original condition. A change in transparency is most naturally attributed to a change in the molecular pattern of the protein constituents of the muscle.

(3) In tortoise muscles (Hill 1950a) the heat production in a twitch begins considerably earlier than the mechanical response; the same is probably true of frog and toad muscles (Hill 1949d), but these are too rapid to allow the priority of the heat to be established with certainty.

It was decided, therefore, to apply the method of quick stretches to a tortoise muscle at 0° C, the speed of which is only about one-tenth to one-twentieth of that of a frog's muscle at the same temperature, so making much smaller demands on the speed of the equipment used in applying the stretch. A Levin-Wyman ergometer was used, with modifications recently described (Hill 1950b), allowing the tension to be recorded against time on a cathode-ray tube, before, during and after a stretch. The stretch started at any desired moment after a shock, was of any desired magnitude, and could be set to any speed required up to 80 mm./sec. Since the stretch found necessary was of the order of 3 mm. it could be completed within 40 msec. if needed; the latent period of tortoise muscle at 0° C being 60 to 90 msec. this gave plenty of scope for examining the mechanical condition of the muscle within the latent period itself.

The experiments were not as simple as had been hoped from previous experience with frog and toad muscles. In those, a rapid stretch applied to a resting muscle (to not too great a length) gave very little rise of tension, so that the tension set up on stretching after a shock could be attributed solely to the active state by then developed. In tortoise muscles at rest, at any length at which they are not actually slack, the tension rises sharply during a stretch (figures 1, 2 and 6), stops rising directly the stretch stops, and then falls gradually to the final value depending on the length. This does not imply that a stretch 'stimulates' the muscle; for the response to a stimulus (figure 3) cannot be detected for 60 to 90 msec. and reaches

a maximum only in 3 to 5 sec., whereas the tension resulting from a stretch is immediate.

This interesting effect, which is discussed later in this paper, provides a complication in the present experiments. In order to avoid it a comparison had to be made between the tension records of a stretch with and without a preceding stimulus. In figure 1 the lower curve of each pair is the tension reached in a muscle stretched 3 mm. in 60 msec., but unstimulated; the upper curve of each pair is the tension recorded in the same muscle similarly stretched, but stimulated 0 or 13 msec.
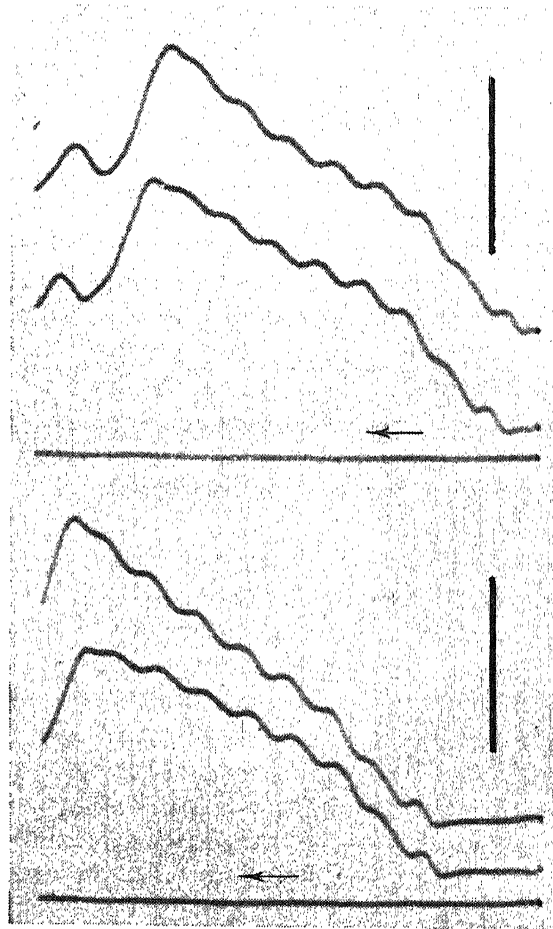
FIGURE 1. Tension developed during a 3 mm. rapid stretch (in about 60 msec.) of a tortoise iliotibialis (500 mg., 35 mm. long) at 0° C. Single sweeps, right to left, 80 msec. Two double records, untouched; in each the lower curve (taken first) refers to a stretch without preceding stimulus, the upper curve to a stretch preceded by a maximal shock at zero time on the right. In the upper record of the lower pair the stretch began 13 msec. after the shock, of the upper pair almost exactly at the same moment as the shock. Note that the lower curve of either pair (i.e. without stimulus) was not necessarily timed to start exactly at the same moment as the upper; the exact comparison is made as shown in figure 2. The vibrations on the curves are due to the sudden release of the armature of the magnet and the high acceleration of the recording equipment at the start of the stretch; they repeat themselves exactly and provide no difficulty in the comparison. The bars on the right denote 50 g. tension.

before the stretch began. The two curves of each pair begin to diverge at about 30 msec., which is about half the latent period determined in the usual way at the greater length (figures 3 and 4). Whereas in the muscle carrying out an ordinary isometric contraction the tension had not risen by 0·1 g. at 60 msec., and had reached only 1·6 g. at 100 msec., the upper curves of figure 1 had diverged from the lower by 10 g. at 60 msec.
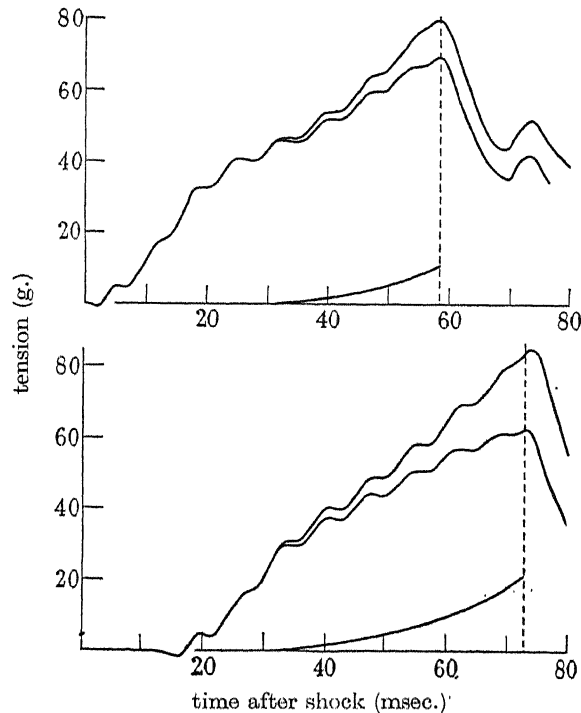
FIGURE 2. Tension developed during a 3 mm. rapid stretch (in about 60 msec.) of a tortoise iliotibialis at 0° C. Upper curve of each pair, stretch following maximal stimulus at zero time; lower curve, stretch without stimulus. Copied in an enlarging lantern from the records of figure 1 and superimposed. The stretch in the lower pair started 13 msec. after the shock and ended at 73 msec., as shown by the vertical broken line; in the upper pair it started at 0 msec. and ended at 58 msec. The small curves at the bottom represent the difference of tension due to the shock. For further details see the legend of figure 1.

The difference between the upper and lower curves in each pair of figure 1 can be attributed to nothing except the stimulus; everything else was the same, and the records could be repeated at will. In each comparison the stretch without stimulus was applied first, to make sure that any possible injury by the stretch did not contribute to the result. The effect is shown best in a short muscle in which little time is spent in the propagation of contraction to the ends of the fibres. The M. extensor iliotibialis (25 to 30 mm.) shows it better than the M. iliofibularis (60 to 70 mm.; formerly described as the biceps cruris). A maximal stimulus was applied, a condenser discharge of time constant about 10 msec. and peak voltage about 8V, between platinum electrodes about 5 mm. apart near the middle of the muscle, which was in oxygen. In order to ensure

that there was no slack in the muscle when the stretch (with or without preceding stimulus) was applied, the following procedure was adopted. For a 3 mm. stretch the muscle was first released 5 mm., then stimulated with two or three shocks to make it shorten up. When it had fully relaxed it was gently drawn out 2 mm. to make sure that it was taut. Finally, the stretch, with or without a preceding stimulus, was applied. This procedure is rather important, for unreliable records may be obtained if there be an unknown amount of slack caused by the release in preparation for a subsequent stretch. It is similar to that previously described (Hill 1949f, pp. 424, 429).

The records of each pair in figure 1 were copied in a lantern on to squared paper and adjusted to coincide at the start; this adjustment could be very exact. Figure 2 shows the comparison. The small curve below each pair represents the difference between the two records. The mean value of the difference in four similar pairs is shown by the broken line of figure 3, for comparison with the much slower development of tension in an ordinary isometric contraction. In figure 4, from another muscle, a similar comparison is given. It is clear that the active state of resistance to stretch is highly developed long before any ordinary sign of contraction appears.

In a recent paper (1950a) it was shown that in tortoise muscle at 0° C the heat production started about 60 msec. after a shock, while the mechanical response could not be detected before about 100 msec. The tension reached its maximum in a twitch in 4 to 5 sec. The experiments had been made during the summer on the muscles of animals kept in a garden. In the present experiments, made during the winter, the tension reached its maximum in 3 to 4 sec., and the latent period was 60 to 70 msec. In both series, the latent period was about 2 % of the time to maximum. In the earlier experiments the heat was observed to start at about 60 % of the latent period after a shock. In the later experiments the state of diminished extensibility could be detected at about 50 % of the latent period. The early 'latency relaxation' and the early change of transparency begin (in frogs' muscles) at about 50 % of the latent period. Putting all these together it is natural to suggest that the 'latency relaxation' and the change of transparency are products of the considerable alteration of mechanical state made evident by a quick stretch, and that the onset of this mechanical change is simultaneous with the start of the heat production.

The heat production at this early stage is presumably heat of activation (Hill 1949a, d); no detectable shortening occurs till later, so that heat of shortening cannot have contributed to the earliest observed phase of the heat. In a previous paper (1950a), it was concluded that the priority of the heat over the mechanical response 'made it impossible to accept, at least in its simple form, the theory that the mechanical energy for contraction is derived from some reservoir or accumulator where it remains in a latent form until released by a stimulus'. The present experiments provide a complication in the issue. If we could regard the sudden increase of rigidity made evident by the quick stretch as a sign of mechanical energy available for shortening and performance of work, the conclusion would be invalid. We know that shortening as such is accompanied by extra heat (Hill 1949a). Can we regard the external work, which has no corresponding quota of heat (Hill 1949c),
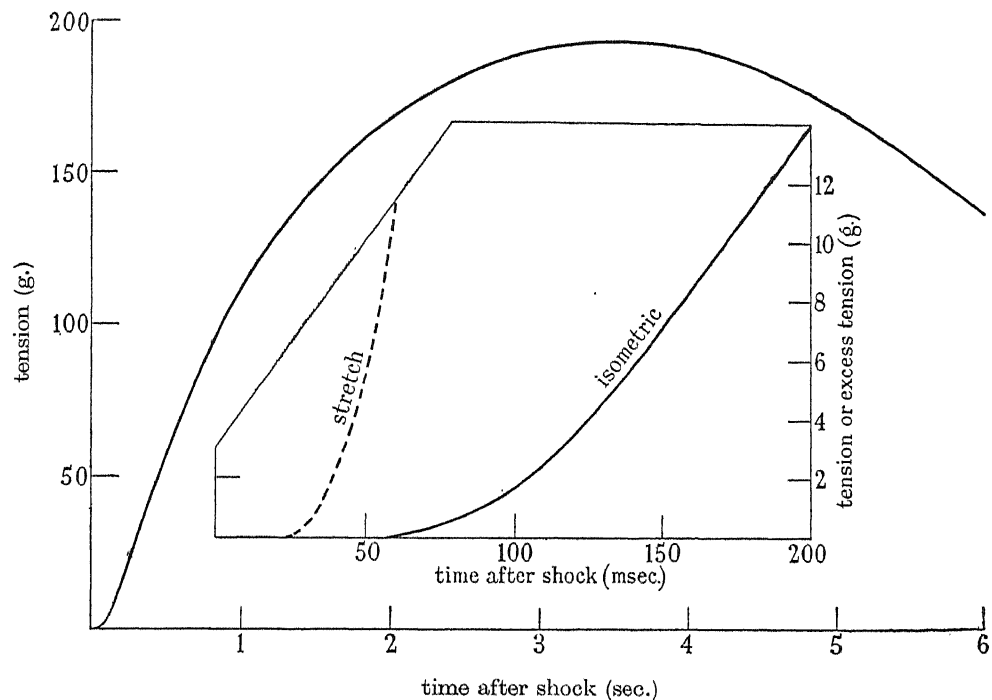
FIGURE 3. Maximal isometric twitch of tortoise iliotibialis at 0° C, at the length to which it was stretched in the records of figures 1 and 2. Main curve, up to 6 sec. Inset, curve up to 0·2 sec. to show latent period (about 60 msec.). Dotted curve to the left (to show the early onset of the active state), the mean of four difference curves similar to those as the bottom of each half of figure 2.
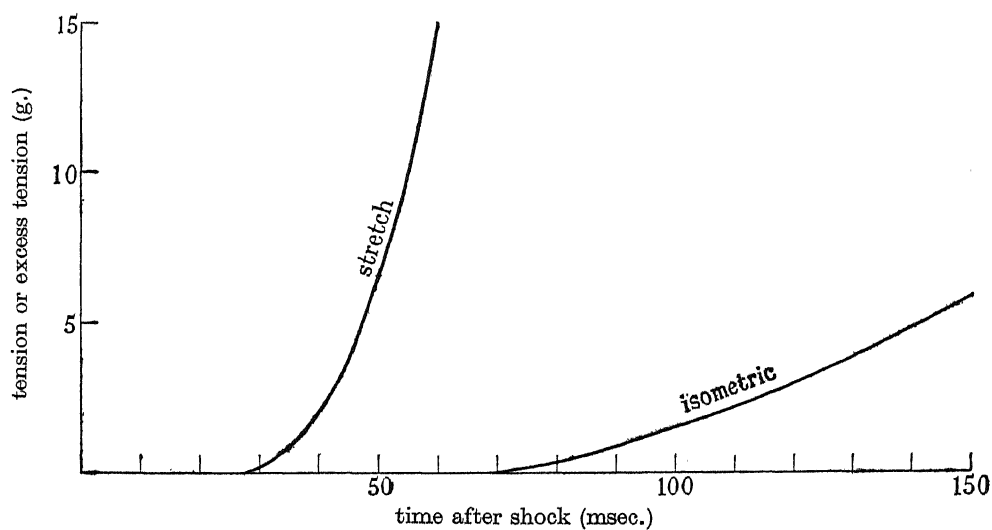
FIGURE 4. The development of the active state after a shock. Iliotibialis of tortoise, 0° C. Right, tension in isometric twitch. Left, excess tension (tension in stretch preceded by shock minus tension in stretch without shock); mean of four pairs of records. 3 mm. stretch in about 55 msec. starting about 6 msec. after a maximal shock. Maximum tension in isometric twitch 192 g.

as being derived from the enhanced rigidity made evident by the stretch? If so, when a muscle shortens against zero load, doing no work, and then relaxes, the mechanical energy potentially available should appear as heat. It does not, there is no extra heat when a muscle relaxes under zero load (Hill 1949b). It seems unlikely, therefore, that the observed rigidity can be regarded as a state possessing mechanical energy potentially available for external work. If so, the argument might still stand but it becomes rather involved. One is tempted, therefore, to conclude that such theories have really outlived their usefulness, that the facts— which are clear enough—can more profitably be discussed in quite different and more specific terms.

It was suggested above that the 'latency relaxation' is a product of the change of rigidity detected by a stretch. The 'latency relaxation' is a very small affair, detectable because of the extremely high sensitivity available in the piezo-electric method. The fall of tension amounts to only about 1 part in a 1000 of the tension later developed during contraction. The curves of figures 1, 2, 3 and 4 show an altogether different order of quantities for the extra force, due to the onset of activity, detected by a stretch; in fact, as figures 6 and 7 show, a large part of the whole tension developed in an isometric twitch can be held quite early by a muscle stretched just after a shock. It seems that the onset of rigidity is accompanied by a slight lengthening, of the order of a few microns, which in a muscle under standing tension causes a drop of tension capable of being detected by very sensitive means. This lengthening is much too small to affect the results of the present experiments. It is presumably an accompaniment, like the change of transparency, of the molecular (or electrical) rearrangement of which the change of rigidity is a sign.

Figure 5 and figure 6 D show the fall of tension following the rise due to stretch of an unstimulated muscle. In the iliotibialis of the tortoise this fall to the final level characteristic of the length to which the muscle has been stretched occupies a few seconds. The time taken in contraction and relaxation following a stimulus is greatly reduced by a rise of temperature; the isometric myogram at 20° is exactly similar to that of figure 3, but on one-eleventh or one-twelfth of the scale of time. It had been expected that the fall of tension after the rise due to stretch of an unstimulated muscle would be similarly quickened by a rise of temperature. It is not; a few preliminary experiments showed little effect of temperature. The small-ness of such an effect is striking and should be further investigated. It means presumably that the energy of activation is small in the molecular rearrangement accompanying the fall of tension. It is mentioned here as evidence, if more is needed, that the fall of tension following the stretch of an unstimulated muscle has little in common with the fall of tension in relaxation after contraction, which has a very high temperature coefficient.

In a previous paper (1949e) the effect was studied of stretching a frog or toad sartorius at various moments during contraction. In these muscles the phenomena are uncomplicated by any significant rise of tension due to stretching without a preceding stimulus, and they are not so rapid at 0° C that the later phases of the change of mechanical state cannot be adequately analyzed by rapid stretches timed to occur at various moments after a shock. The complication in tortoise muscle,

due to changes of tension caused by stretch without stimulus, cannot be altogether avoided by subtracting tension without stimulus from tension following stimulus, since one does not know whether the former occurs in the same structure as the latter or in a parallel but independent one.
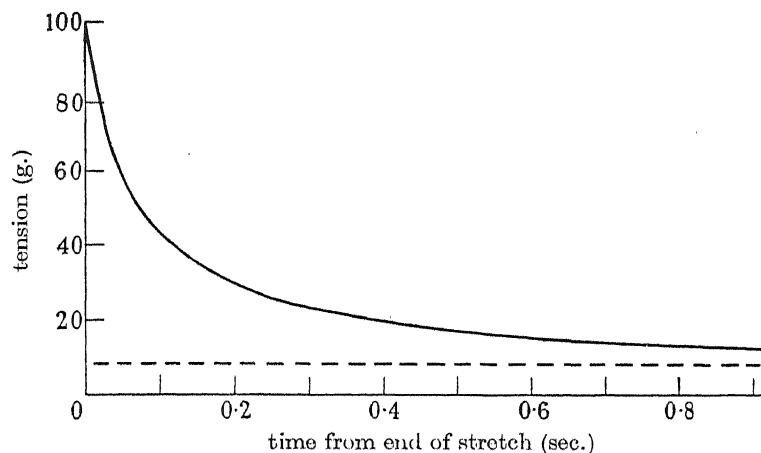
FIGURE 5. Fall of the tension set up by a rapid stretch (5 mm. in 80 msec.) of an unstimulated tortoise iliotibialis at 0° C. The stretched length was the same as that of the records in figures 1, 2 and 3. The lower broken line shows the nearly constant level (about 8 g.) reached after about 10 sec.
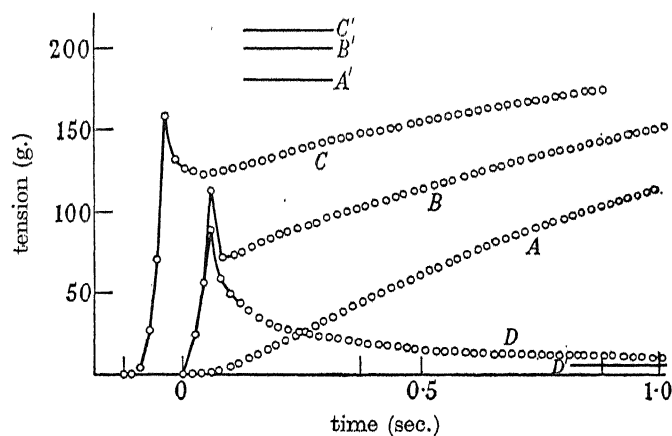
FIGURE 6. Effect of stretching a muscle on the tension exerted during a twitch. Iliotibialis of tortoise at 0° C. The recording beam was modulated to give 50 spots per sec., and six 1 sec. sweeps were photographed covering contraction and early relaxation. To avoid confusion in the figure only the first 1 sec. of each contraction is shown, together with the height of its maximum. The records, placed in a lantern, were copied on to squared paper on the same base line: for clarity, the early spots are joined by lines. *A*, isometric twitch, maximum at *A'* at 3·2 sec. *B*, 4mm. stretch starting at the same moment as the shock and taking 60 msec. Maximum at *B'* at 2·8 sec. *D*, 4 mm. stretch without stimulus, taking 60 msec. Tension after 4 sec. at *D'*. *A*, *B* and *D* are plotted from the same origin (shock or beginning of stretch). *C*, 4 mm. stretch starting 60 msec. after the shock and taking 60 msec. Maximum at *C'* at 2·6 sec. Origin (shock) 0·14 sec. to the left. The final length of *B*, *C* and *D* was the same as that of the isometric twitch *A*. The latent period in an isometric contraction was about 62 msec.

Apart, however, from this complication, experiments show that the after-effect of stretch, and the effect of stretch at later moments, are entirely similar to those met in frog and toad muscle. Figure 6, for example, shows the effect of 4 mm. stretch timed to occur at or just after a shock, for comparison with the effect of a stretch without shock or a shock without stretch. Figure 7 shows the effect of three stretches at 16, 50 and 150 msec. after a shock, for comparison with an isometric contraction. The curves of figures 6 and 7 bear an obvious resemblance to those of figures 3 and 4 of the previous paper. They are useful in confirming the very early development of the active state, for the times at which the stretches started in figures 6 and 7 are very short compared with the time to maximum in an isometric twitch, e.g. even for curve D, figure 7, 150 msec. as compared with 4 sec., about 4 %. With stretches of rather greater extent the tension curves run horizontal for a time (until relaxation sets in) just like those of figures 1 and 2 of a previous paper (Hill 1949e).
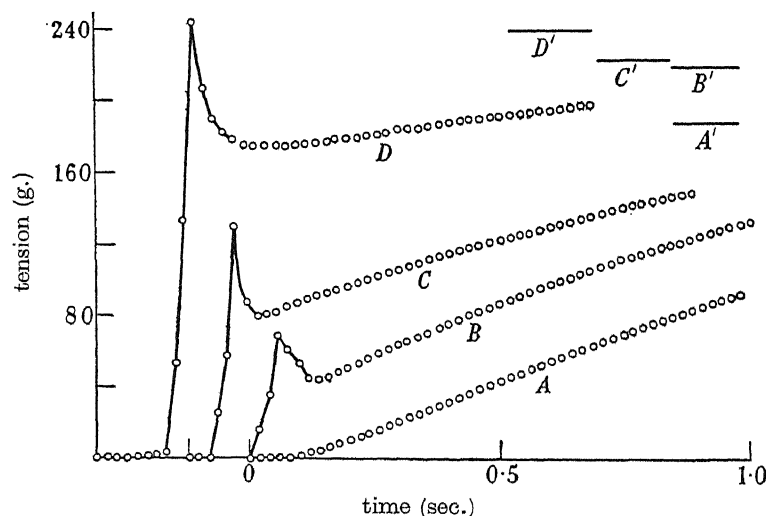
FIGURE 7. Effect of stretching a muscle on the tension exerted during a twitch. Iliotibialis of tortoise at 0° C, the opposite muscle of the same animal as used for figures 1, 2 and 3. For details see legend of figure 6. A, isometric twitch, maximum at A' at 4 sec. B, 4 mm. stretch starting 16 msec. after the shock and taking 60 msec. The final (stretched) length of B was the same as that of the isometric twitch A. Maximum at B' at 4·4 sec. A and B are plotted from the same origin (the shock). C, 4 mm. stretch in 60 msec. starting 50 msec. after the shock. Final length the same as A. Maximum at C' at 4·3 sec. Origin (shock) about 120 msec. to the left of that of A and B. D, 4 mm. stretch in 60 msec. starting 150 msec. after the shock. Final length the same as A. Maximum at D' at 3·6 sec. Origin (shock) about 190 msec. to the left of that of C. The records from which figure 4 was constructed were made immediately before those given here. Latent period about 70 msec.

It appears, therefore, that after a stimulus, about half-way through the latent period as ordinarily measured, a state of enhanced rigidity starts to develop. Its onset is very rapid, its full extent being reached in a time which is a small multiple of the latent period, only a few hundredths of the time taken by an isometric twitch to reach its maximum. As relaxation sets in the condition of enhanced

rigidity gradually disappears. The form of the twitch as ordinarily recorded depends practically not at all on the speed of onset of the active state, but only on the velocity of shortening as a function of load, as defined by the characteristic relation between force and speed. The maximum in an isometric twitch is reached as a balance between continuing shortening and oncoming relaxation and occurs after relaxation has already begun, as is shown by the fact that the decline of tension after stretch begins earlier than in an isometric twitch.

We may conclude, therefore, that the fundamental result of stimulation is a rapid change of mechanical condition which reaches its climax very early, lasts for a time and then relaxes slowly. The two chief properties of this altered mechanical state are: (1) a large increase in rigidity, i.e. in resistance to stretch, and (2) a capacity to shorten with a velocity depending on the characteristic relation between force and speed. The 'latency relaxation' and the early change of transparency are physical signs of the altered mechanical condition, presumably of an altered molecular pattern of the muscle proteins; while the heat of activation can be regarded as a sign of the chemical changes by which the mechanical and physical alterations are brought about. As shortening occurs a proportional amount of heat is set free, perhaps in reactivating a system the activity of which has been dissipated by shortening.

REFERENCES

Abbott, B. C. & Ritchie, J. M. 1948 *J. Physiol.* **107**, 5*P*.
Abbott, B. C. & Ritchie, J. M. 1950 *J. Physiol.* (in the Press).
Hill, A. V. 1949*a* *Proc. Roy. Soc.* B, **136**, 195–211.
Hill, A. V. 1949*b* *Proc. Roy. Soc.* B, **136**, 211–219.
Hill, A. V. 1949*c* *Proc. Roy. Soc.* B, **136**, 220–228.
Hill, A. V. 1949*d* *Proc. Roy. Soc.* B, **136**, 242–254.
Hill, A. V. 1949*e* *Proc. Roy. Soc.* B, **136**, 399–420.
Hill, A. V. 1949*f* *Proc. Roy. Soc.* B, **136**, 420–435.
Hill, A. V. 1950*a* *Proc. Roy. Soc.* B, **137**, 268–273.
Hill, A. V. 1950*b* *Proc. Roy. Soc.* B, **137**, 273–280.
Hill, D. K. 1949 *J. Physiol.* **108**, 292–302.
von Muralt, A. 1934 *Pflüg. Arch. ges. Physiol.* **234**, 653–664.
Sandow, A. 1944 *J. Cell. Comp. Physiol.* **24**, 221–256.
Sandow, A. 1947 *Ann. N.Y. Acad. Sci.* **47**, 895–929.

A note on the heat of activation in a muscle twitch

By A. V. HILL, F.R.S.

Biophysics Research Unit, University College, London

Under extreme loads a stimulated muscle neither shortens nor develops tension. The heat production in a twitch is then rather less than one-half of its value with maximal shortening. Under such conditions the heat of shortening and the mechanical work are nil. The remaining heat, therefore, is heat of activation alone. It is about the same in magnitude and onset as the heat of activation at ordinary lengths.

The energy liberated in a muscle twitch appears in three forms (Hill 1949, *a*, *b*), heat of activation, heat of shortening and mechanical work. The separate identity of the activation heat could best be demonstrated if shortening could be altogether prevented; for then heat of shortening and work would both be absent. The ordinary isometric twitch is of no use for this purpose, since substantial internal shortening occurs, and considerable work is done which later turns into heat in relaxation. By stimulating a muscle several times under zero load it can be made to draw up so far that not much further shortening is possible in a subsequent contraction and little but heat of activation is then left; this, however, is not entirely satisfactory, partly because the muscle may not lie close enough to the thermopile under zero tension but also because some of the fibres in the shortened state may be slack and so actually able to shorten by taking up slack when stimulated. In an isometric tetanus, once a constant tension has been developed no further shortening occurs nor is work done, and the activation heat now appears alone, in a summated form, as heat of maintenance.

As the initial load, or tension, on a muscle is increased the amount of shortening in an isotonic contraction, and the extra tension developed in an isometric contraction, diminish; at a high enough load or tension both become zero. This is most strikingly seen in the work of Ramsey & Street (1940), who showed that single muscle fibres can be stretched without rupture as much as 100 %, and that the extra tension developed on stimulation is then nil. The same general effect, however, was found by many previous workers. Evans & Hill (1914) showed that in isometric twitches the extra tension developed became very small at great lengths and the heat production was substantially reduced. Doi (1920, 1921) obtained similar results. Hill (1925) at an extension of 40 % found the extra tension developed reduced to less than 20 % and the heat to about 70 %. The only difficulty in such experiments is the danger of rupture. If this is avoided by careful dissection and attachment repeatable results are obtained, apart from a certain amount of permanent stretch.

Simultaneous records were made of heat and mechanical response with various loads, from very small to very large. At small loads the shortening heat is large and the work small; at medium loads the shortening heat is less and the work

greater. At great loads the shortening heat becomes negligible but the work (being the product of small shortening by large load) may remain finite, and this provides a complication since it turns into heat in relaxation. Only at extreme loads does the work become negligible. The record of an isotonic contraction shows when this state is reached; it requires a load of several kg./cm.² in a good muscle.

The result is as follows. When all shortening is apparently abolished the heat production that remains is rather less than one-half the total heat given out with very small loads. This is consistent with the previous conclusion (Hill 1949a) that the heat of activation is rather less than the maximum heat of shortening. It starts off very rapidly soon after the stimulus, in fact, its initial course is scarcely distinguishable from that of the heat in an ordinary contraction; only when the heat of shortening begins to make a substantial contribution do the two clearly diverge. One has the impression that the heat production under extreme load is rather more protracted than the activation heat determined by subtracting the heat of shortening in an ordinary contraction. This would be in keeping with the fact (Hartree & Hill 1921) that the twitch itself is more drawn out at greater extensions. But it might be due partly to a small amount of work being done internally (e.g. one part of a fibre or fibre bundle stretching another part) and turned into heat in relaxation. As a pretty good first approximation, however, the separate identity of the heat of activation can be well exhibited in this way.

It must be admitted that the activation heat so determined is not necessarily the same as at normal lengths. So far, however, as the evidence goes the activation heat is little, if at all, affected by length. There are obvious reasons why heat of shortening and work should be affected by length and load, but it would be easily intelligible that activation heat should not be so affected, if we regard it as an accompaniment of the chemical processes required to induce the altered state of activity. The heat of maintenance during a tetanus does depend somewhat on length (Fenn & Latchford 1933; Abbott 1950), but this dependence is probably associated with the fact that relaxation occurs more slowly at greater lengths. In fact, if activation heat is more protracted at a greater length its summated effect per unit of time in a maintained contraction should be less.

REFERENCES

Abbott, B. C. 1950 *J. Physiol.* (in the Press).
Doi, Y. 1920 *J. Physiol.* **54**, 218–226.
Doi, Y. 1921 *J. Physiol.* **55**, 38–42.
Evans, C. L. & Hill, A. V. 1914 *J. Physiol.* **49**, 10–16.
Fenn, W. O. & Latchford, W. B. 1933 *J. Physiol.* **80**, 213–219.
Hartree, W. & Hill, A. V. 1921 *J. Physiol.* **55**, 389–411.
Hill, A. V. 1925 *J. Physiol.* **60**, 237–263.
Hill, A. V. 1949a *Proc. Roy. Soc.* B, **136**, 195–211.
Hill, A. V. 1949b *Proc. Roy. Soc.* B, **136**, 220–228.
Ramsey, R. W. & Street, S. F. 1940 *J. Cell. Comp. Physiol.* **15**, 11–34.

The application of phase-contrast to the ultra-violet microscope

By E. W. Taylor, *Cooke, Troughton and Simms, Ltd., York*

[Plates 15 to 17]

It has been shown that the phase-contrast principle of Zernike may be applied with advantage to ultra-violet microscopy and that certain advantages follow:

(*a*) In the first place, it furnishes a more certain method than dark-ground illumination for the visual selection of a suitable field of view as all structures are more clearly revealed.

(*b*) In the second place, it furnishes a new means whereby structural details that give rise to a change of phase in the transmitted radiations may be photographed in contrast, even under conditions when selective absorption does not take place.

(*c*) In the third place, it is possible, after having selected a suitable field by means of visual phase-contrast, to photograph the same field in ultra-violet light and then to turn once again to the visual phase-contrast image to make certain that the exposure has not caused damage to a living organism.

A series of explanatory photographs has been taken covering a wide range of biological objects. These show that the method is sensitive to minute changes in phase and that the resulting images are characterized by good contrast. Details are revealed that cannot be brought out by normal methods, as will be clear from a close study of the pairs of photographs which accompany this paper. In selecting these it was thought necessary to avoid very fine detail and delicate shades of contrast that would be lost in the process of reproduction.

INTRODUCTION

The phase-contrast microscope of Zernike is now widely used for studying the structure of those transparent objects which produce small changes of phase in the transmitted light. As is well known, such changes are transformed into changes of amplitude and give rise to a contrasted image in which the brightness at any image point is related to the phase change produced at the corresponding object point.

The principle of interference, by means of which a change of phase is transformed into a change of amplitude, is not, of course, confined to those wave-lengths to which the human eye is sensitive. If, for instance, an object illuminated by ultra-violet radiation gives rise to changes of phase in the transmitted light, these may also be converted into changes of amplitude which, though invisible to the human eye, may be readily recorded on the photographic plate.

The object of this paper is to describe apparatus designed for this purpose and to illustrate some of the practical results obtained. Reference is also made to the value of visual phase-contrast in selecting a suitable field of view for subsequent photography in the ultra-violet region of the spectrum.

HISTORICAL

The practical realization of the use of ultra-violet radiations in the microscope was made possible when von Rohr computed objectives which were corrected for the wave-length 2750 Å of the cadmium spark. Kohler (1904) designed the apparatus

with which these lenses could be used and demonstrated that the ultra-violet microscope possessed about twice the resolution normally obtained by a visual instrument equipped with objectives of similar numerical aperture. He also showed that considerable variation existed in the absorption of structures present in certain biological objects and pointed out in addition that ultra-violet radiations often brought about pronounced physiological changes in living cells.

The early use of this equipment presented a number of difficulties. The electrical apparatus for producing a condensed spark was crude by present-day standards, and consequently the output of radiant energy was low. The method of focusing by means of a searcher eyepiece was unsatisfactory, and the fine focusing adjustment of the best Zeiss microscope of that day was not sufficiently sensitive or accurate for the purpose. (The focal depth of ultra-violet objectives is about half that of the corresponding visual lenses and for the highest powers is only about 0.25μ.)

Barnard in 1924 (Barnard & Welch 1936) had made a duplex illuminator in which a Zeiss quartz condenser was mounted co-axially in a dark-ground system—the front lenses of the complete condenser being common to both illuminators. When the quartz condenser was stopped out the dark-ground system could be used with visual light in combination with a special visual objective, designed to work through a quartz cover-glass. By this means objects could be examined and fields selected for later study by means of ultra-violet microscopy. To facilitate this procedure the visual focus of the dark-ground system was made to coincide with that of the quartz condenser by means of suitable adjustments.

Barnard also experienced the difficulty of precise focusing, due to the insensitivity of the fine adjustment, and this was overcome when the Beck-Barnard microscope was produced in 1926. At a later date a special objective changer, by means of which the visual and ultra-violet objectives could be centred to one another, was developed by Welch. This was later modified by Barnard to the more compact form still in use, and the parts were so dimensioned that all objectives were nearly parfocal, the residual error being determined and allowed for by adjustment of the fine focusing mechanism.

PHASE-CONTRAST IN VISUAL LIGHT

The phase-contrast microscope of Zernike is now widely used for studying the structure of those transparent objects which produce small changes of phase in the transmitted light. As is well known, such changes are transformed into changes of amplitude and give rise to a contrasted image in which the brightness at any image point is related to the phase change produced at the corresponding object point. Contrast is at a maximum for those structures which give rise to diffraction spectra in which the total combined amplitude is equal to that of the null order transmitted by the absorbing annulus.

For the condition of maximum contrast there is greater point-to-point correspondence between object and image than is the case when ordinary illumination is used and when the condenser aperture must be reduced to enhance the otherwise weak contrast. By the latter method the image is degraded by diffraction fringes

and is not an adequate representation of the object. Resolution is also lost, due to restriction of objective aperture, whereas by phase-contrast almost the full aperture of the objective is used to form the image and the resolution falls little short of that to be expected when an amplitude object is illuminated in the ordinary manner and the objective is used at full aperture.

In relatively few instances are cell structures not revealed by phase-contrast. In most cases those which have been disclosed by cyto-chemical methods are revealed with great clarity. Such structures can, however, only be differentiated according to their size and form and cyto-chemical methods remain as the only means by which differentiation according to chemical constitution can be carried out

It may be pertinent here to quote from the *Twenty-fifth Annual Report of the British Empire Cancer Campaign*: 'The differences in refractive index between the various structural components of living cells are insufficient to render them distinct with the ordinary microscope, since they also lack specific absorptions in visual light. These difficulties have been overcome by the introduction of the phase-contrast technique, which accentuates differences of refractive index. By this method, cells of a variety of different tumours have been studied. Mitochondria and cytoplasmic granules are revealed with the greatest clarity, as are also the chromosomes of dividing cells. Better opportunity is thus afforded for studying in the living condition the cellular organellae which previously could only be adequately demonstrated by employing the more elaborate cytological techniques.'

The apparatus commercially available for the phase-contrast technique has been designed for visual purposes and the aim has been to obtain the best possible image contrast at those wave-lengths for which the human eye is most sensitive.

PHASE-CONTRAST IN ULTRA-VIOLET LIGHT

It is, of course, possible to apply the method to the examination of objects illuminated by ultra-violet light of wave-lengths to which the eye is insensitive, and by this means to convert changes in phase, produced by the object, into changes of amplitude produced in the image recorded by the photographic plate. Already an attempt has been made to extend the principle of phase-contrast to the use of the longer wave-lengths in the ultra-violet and Foster has recorded its use in the 3650 Å mercury line using a special objective produced in the United States of America (Foster 1949).

It is particularly desirable to extend the principle of phase-contrast into the ultra-violet region of the spectrum, especially as most biological objects are transparent to the long ultra-violet wave-lengths, and good image formation by normal means cannot be expected for the reasons already given.

Important improvements, as regards contrast, resolution and fidelity in the image, may be expected to follow from the use of the phase-contrast technique, particularly when strong absorption of the radiations by the object does not take place. At shorter wave-lengths, such as 2750 Å, some parts of a living cell absorb strongly, and the contrast in the phase-contrast image may then be due to two distinct causes and the interpretation thereby complicated.

Perhaps at this stage a further extract from the same *Annual Report* may be admitted: 'Also, structures revealed by the ultra-violet light microscope are the result of selective absorptions, which are determined by molecular structure of the absorbing substances. Thus arises the possibility of determining the distribution of certain classes of substances in living cells—a vital cytochemical technique. Direct observation of living cells by ultra-violet light has been found to be impracticable. Sufficient structural detail is not perceptible by employing fluorescent eyepieces. Cytological observations must therefore be made by photographic methods.'

Although objects highly transparent to visible radiations may show considerable contrast, when illuminated by ultra-violet light, owing to selective absorption by the specimen, this is not true at all wave-lengths nor is it to be assumed that all important details of structure can be revealed in this manner. In any case, the contrast due to absorption and that due initially to phase changes produced by the object are unlikely to show similar characteristics in the image, since the two methods make use of independent properties in the object. The study of change of phase may indeed add a new and important technique to those already available to the cytologist.

Selecting the object for ultra-violet microscopy by means of the phase-contrast searcher

When attempting the microscopical examination of objects illuminated by ultra-violet light, one is at once confronted by the difficulty of selecting a suitable field of view before resorting to the photographic plate as a final means of recording the appearance of the otherwise invisible image. For this purpose it has been the practice to make use of a fluorescent screen placed in the focal plane of a viewing eyepiece. This screen, under the influence of the invisible radiations, presents to the eye a visible magnified image of the object, but unfortunately the image so produced is weak, lacking in detail and generally inferior to a normal visual image.

A more satisfactory method is first to examine the object by visual light, in the normal manner, and then to interchange the visual objective for one corrected for the invisible radiations and to do this in such a manner that an identical field of view is photographed. When the object to be examined is highly transparent, the visual image is brightly illuminated but lacking in contrast and under such conditions the contrast can be greatly increased by the application of the Zernike principle. This should be brought about in such a manner, that, having discovered a suitable field of view by means of phase-contrast, it is possible to photograph the identical field of view in the ultra-violet in the normal manner with a minimum disturbance of the instrumental adjustments. In this manner visual phase-contrast may be employed as an aid to those engaged in ultra-violet microscopy, even though the contrast in the final photographic image is due only to selective absorption by the specimen. Dark-ground illumination has been previously used for this purpose, but it has not the penetration of the phase-contrast method and often fails to reveal the chromosomes and other important details.

When using the phase-contrast method in visible light it is customary to examine the back focal plane of the objective by means of an auxiliary microscope which,

on being substituted for the eyepiece, enables the observer to establish coincidence between the image of the condenser diaphragm and the corresponding area of the phase plate. Similarly, when applying the phase-contrast method to the ultra-violet microscope an auxiliary microscope of similar type is used with a quartz objective and a fluorescent screen set in the focal plane of the eyepiece. In this manner a visible image of the phase plate and the condenser diaphragm is formed by fluorescence and a coincidence setting is then made in the normal manner.

The instrumental equipment

A microscope adapted for phase-contrast in the visual and for both phase-contrast and absorption microscopy in the ultra-violet region of the spectrum, requires certain additional parts. They consist of:

(a) An auxiliary condenser lens adapted to compensate for the change in the focus of the quartz condenser as between the visual and ultra-violet radiations.

(b) A removable condenser annulus.

(c) Precise interchangeability between the visual and ultra-violet phase-contrast objectives.

(d) A special auxiliary microscope with a quartz objective and a fluorescent screen.

The optical systems will now be more fully described with reference to the figures 1a to e.

The optical system

(Figures 1 a to e)

(a) *The normal visual system*

An auxiliary condenser A is introduced into the illuminating beam to produce a change in the focus of the quartz condenser C which is normally adjusted for ultra-violet radiations. The object D is then magnified by the objective E to produce a normal visual image in the eyepiece L.

(b) *The visual phase-contrast system*

The condenser annulus B and the visual phase retarding plate G are introduced into the system in conjugate planes and after suitable adjustments have been made a selected portion of the object is imaged in the field of view of the eyepiece L.

(c) *The normal ultra-violet absorption system*

The auxiliary condenser A, the condenser annulus B and the phase plate G are tripped out of the system and the quartz monochromat F is substituted for the visual objective E. The ultra-violet image, with contrast due to absorption, may now be observed by means of the fluorescent screen at K, and subsequently photographed.

(d) *The phase-contrast ultra-violet system*

On inserting into the system the condenser annulus B and the ultra-violet phase plate H a phase-contrast image in ultra-violet light will be formed on the fluorescent screen K and subsequently photographed.
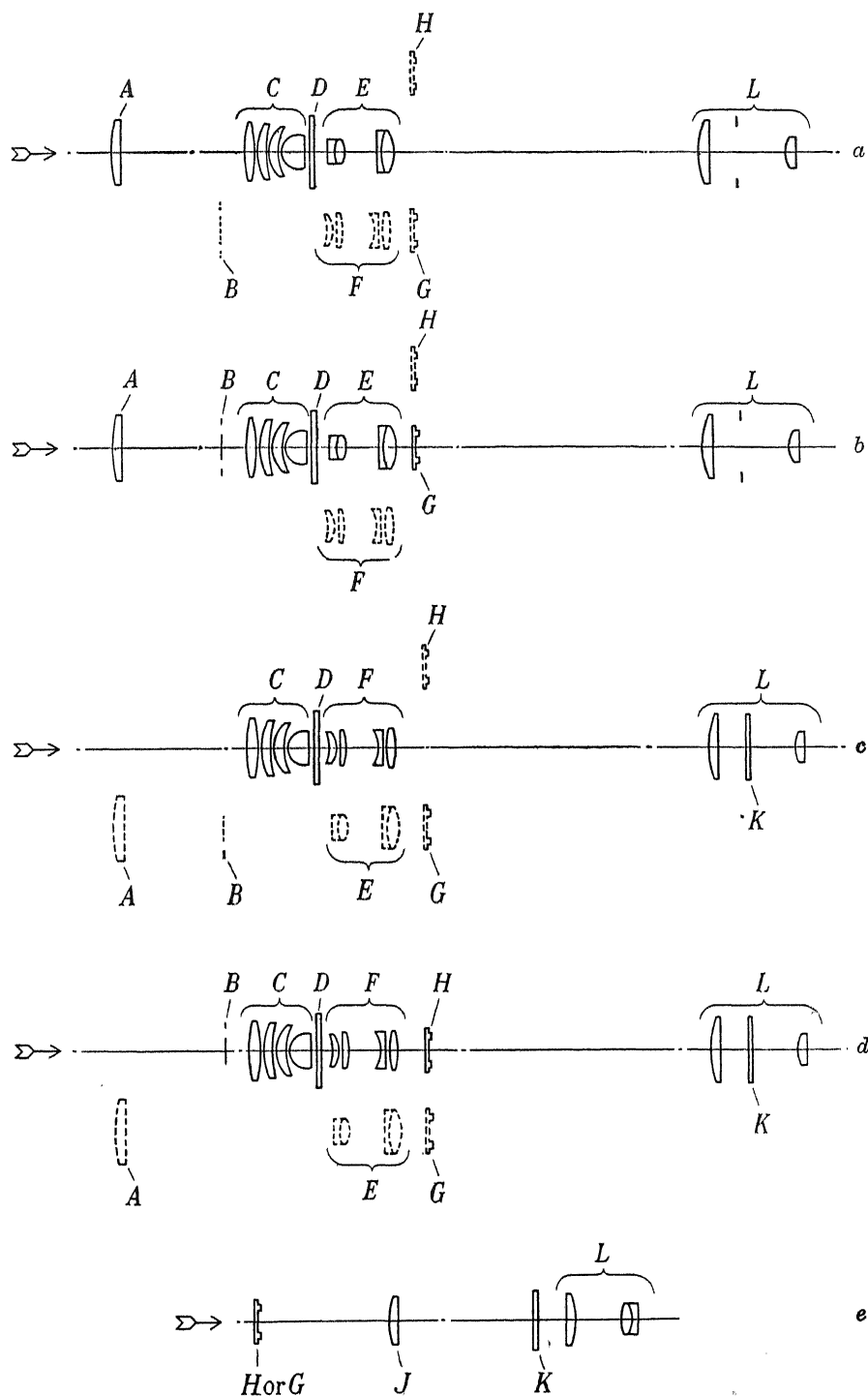
FIGURE 1a to e. For descriptions see text.

Key to lettering

A, auxiliary condenser lens.
B, condenser annulus.
C, quartz ultra-violet condenser.
D, quartz microscope slide.
E, visual objective.
F, ultra-violet monochromat.

G, visual phase plate.
H, ultra-violet phase plate.
J, quartz objective.
K, fluorescent screen.
L, visual eyepiece.

(e) The auxiliary microscope

When using the phase-contrast principle it is necessary to ensure that the image of the condenser annulus formed by the condenser and objective acting together coincides with the phase ring of the phase plates G and H. To achieve this the eyepiece is withdrawn and an auxiliary microscope inserted in its place. The auxiliary microscope consists of a quartz objective J, a fluorescent screen K and a normal eyepiece L. The condenser annulus B is centred by means of adjusting screws until the required coincidence is obtained as judged by the images seen on the fluorescent screen K.

The instrumental adjustments

(1) The axis of the illuminating ultra-violet beam is brought into co-axial relationship with that of the microscope tube.

(2) The axis of the illuminating visual beam is brought into similar relationship.

(3) Using the quartz condenser and a visual objective screwed directly into the microscope body, focus an object, choosing if possible one that absorbs the ultra-violet radiation to be used.

(4) Adjust the quartz condenser to give even illumination over the field of view.

(5) Remove the objective and remount in the adjustable changer. After focusing, centre the objective to the same field of view as seen in (4) by means of the adjusting screws.

(6) Replace the visual with the quartz objective, mounted in a second changer. Illuminate the object with ultra-violet radiation and focus by means of the fluorescent searcher eyepiece.

(7) Centre the quartz objective to the same field of view and readjust the condenser to give even and maximum illumination.

(8) Insert the annulus in the focal plane of the condenser and by means of the quartz auxiliary microscope centre the image of the condenser annulus to the absorbing ring of the 'Z' disk.

(9) Since care has been taken to mount the 'Z' disks centrally in the objectives it will be found on checking the centering of the visual phase-contrast system that reasonably good coincidence exists between the two annuli and that sufficient contrast is present in the image to enable a thorough search to be made of the object.

Taylor

Proc. Roy. Soc. B, volume 137, *plate* 15

2
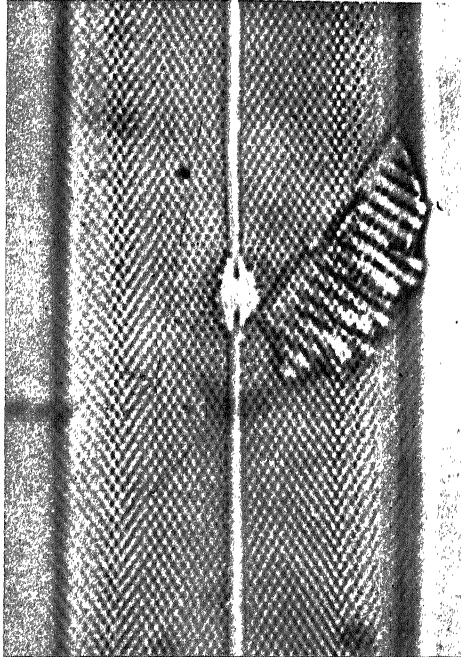
3
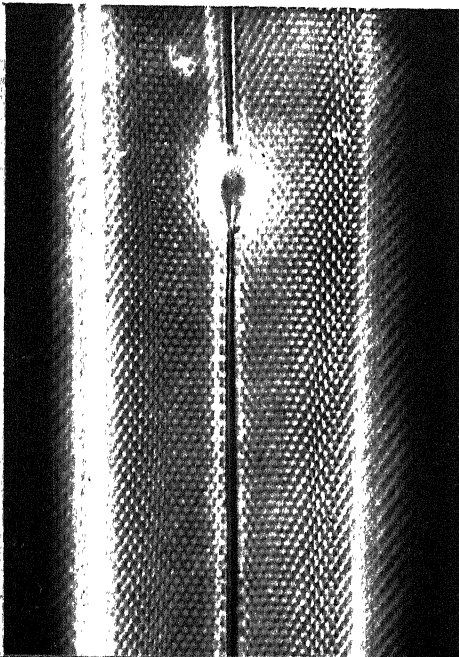
4
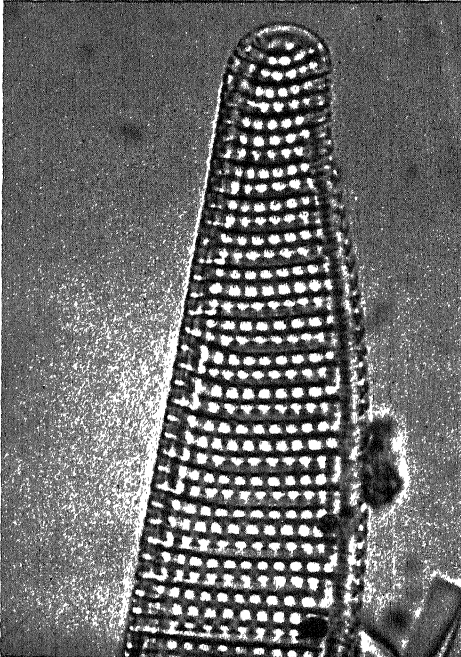
5
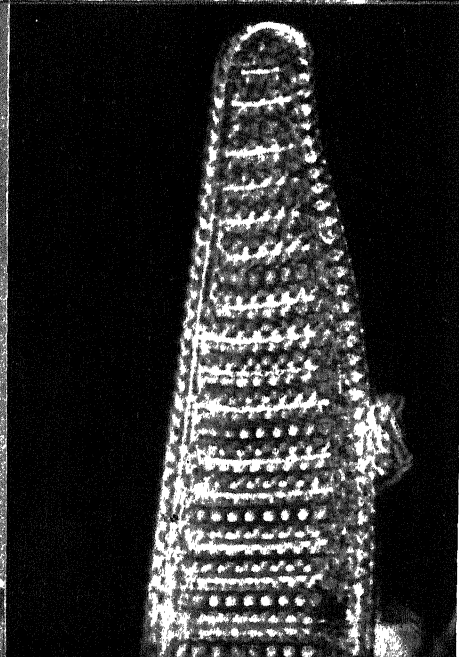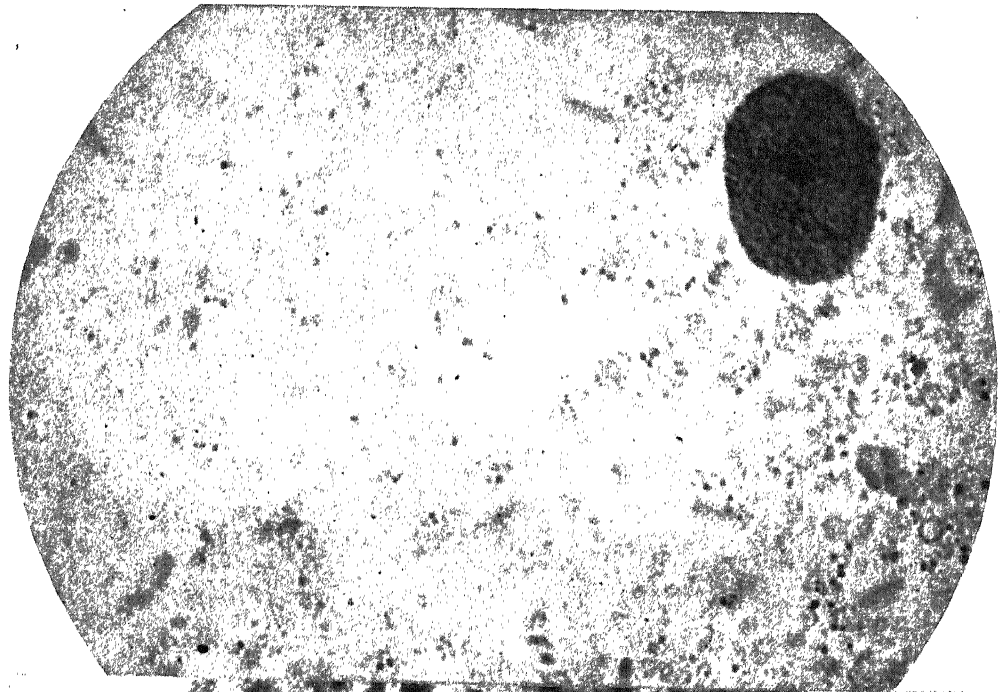
6
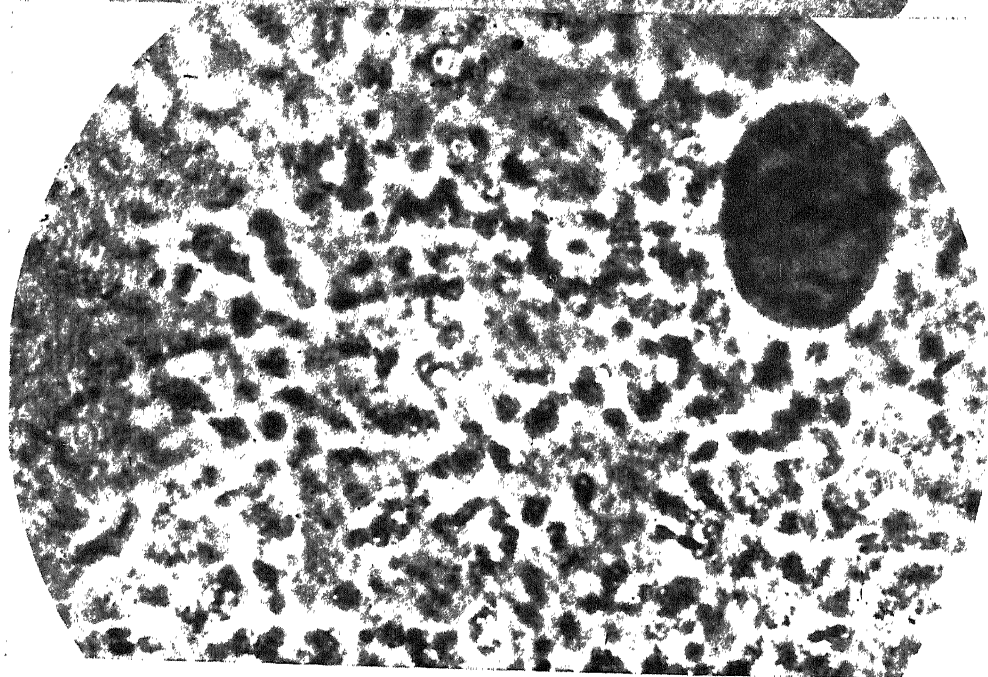
7
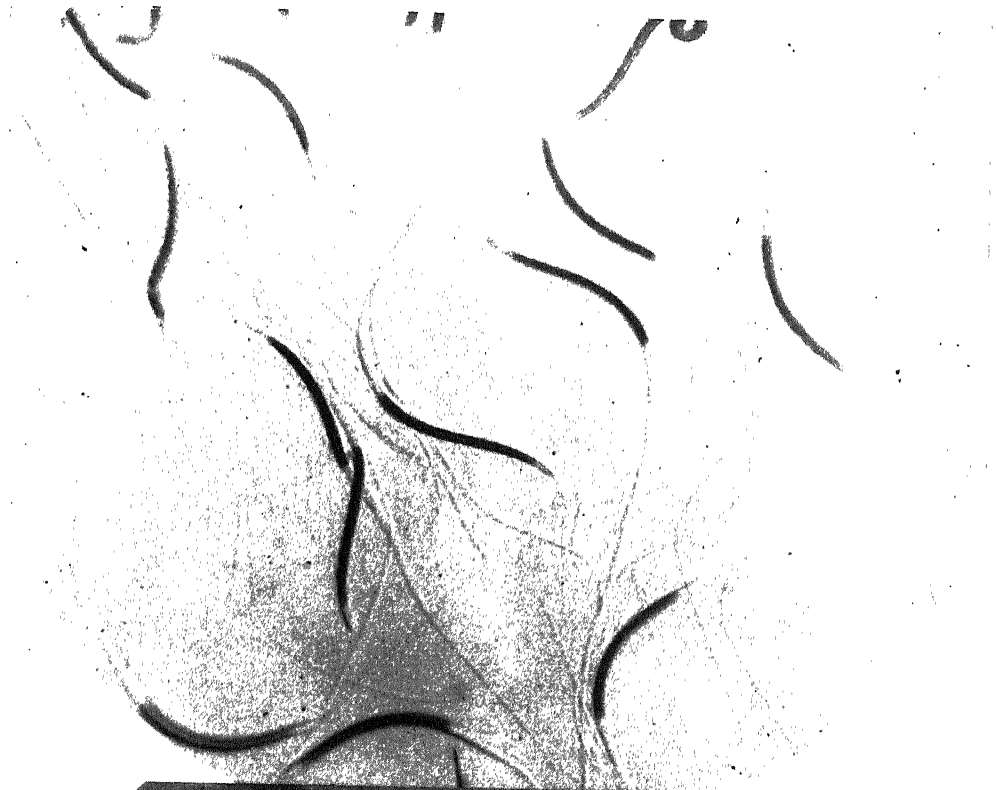
8
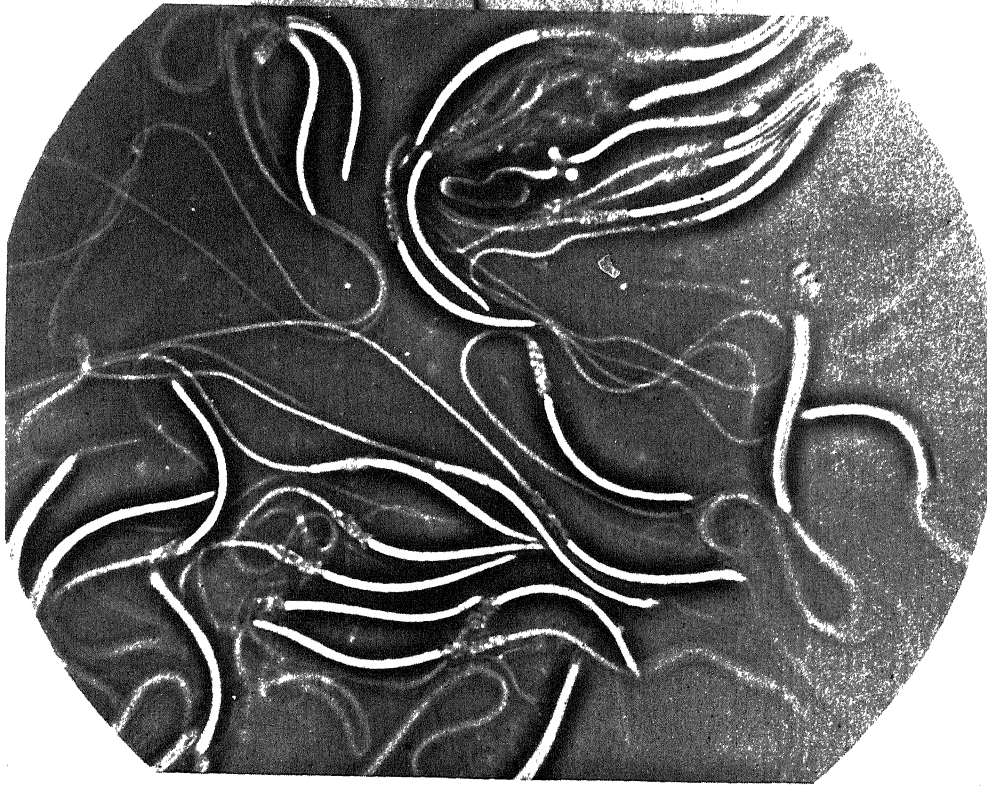
9

REFERENCES

Barnard, J. E. & Welch, F. V. 1936 *Practical photo-micrography*, 3rd ed. London: Arnold.
Foster, L. V. 1949 *Anal. Chem.* **21**, 435.
Kohler, A. 1904 *2. Mikroskop.* **21**, 129, 273.
25th Annual Report, British Empire Cancer Campaign, pp. 64, 65.

DESCRIPTION OF PLATES 15 TO 17

Work still remains to be done in the choice of the most suitable type of plate for this region of the spectrum (2750 Å).

PLATE 15

FIGURE 2. *Pleurosigma delicatalum* (magn. × 2350), transmitted light (N.I.M.R.).
FIGURE 3. *Pleurosigma delicatalum* (magn. × 2350), phase-contrast (N.I.M.R.).
FIGURE 4. *Epithemia turgida* (Ehremb.) Kutz (magn. × 2350), transmitted light (N.I.M.R.).
FIGURE 5. *Epithemia turgida* (Ehremb.) Kutz (magn. × 2350), phase-contrast (N.I.M.R.).

PLATE 16

FIGURE 6. Epithelial cell, transmitted light (magn. × 2500), (C.T. and S.).
FIGURE 7. Epithelial cell, phase-contrast (magn. × 2500), (C.T. and S.).

PLATE 17

FIGURE 8. Living fowl sperm (magn. × 2115), transmitted light. (N.I.M.R.).
FIGURE 9. Living fowl sperm (magn. × 2115), negative phase-contrast. (N.I.M.R.). Note that the middle piece, not differentiated in figure 8, is now clearly resolved to show a spiral formation. Note also the fine tail filaments.

Amoebicidal action and chemical constitution

By J. D. FULTON, L. P. JOYNER, H. KING, F.R.S., J. M. OSBOND
AND JEAN WRIGHT

The National Institute for Medical Research, Mill Hill, N.W. 7

This communication describes some of the results which have been obtained in the search for an amoebicidal agent which might replace emetine. Eighty-seven compounds have been synthesized starting from two 4-chloroquinolines with a methyl or anilino group in the 2-position. A number of compounds have been found with an *in vitro* activity on *Entamoeba histolytica* slightly inferior to that of emetine.

INTRODUCTION

The need for a drug of more general application than emetine became apparent in World War I and was equally pressing during World War II. The search for such a drug should be amenable to experimental methods, but a major difficulty has been and still is the establishment of a reliable biological test of amoebicidal activity.

As early as 1911, tests on the amoebicidal activity of chemical compounds were made by Vedder and a few years later by Pyman & Wenyon (1917), but the results

which they obtained were of doubtful value regarding the treatment of human infections with *Entamoeba histolytica* since the tests were carried out on free living amoebae and, as Dobell & Laidlaw (1926) later showed, the response to emetine in culture varies widely with the species.

In 1925 Boeck & Drbohlav described a method of culturing *E. histolytica* in the presence of bacteria which are essential for growth. They used a biphasic medium with solid base and liquid overlay. This method was improved by Dobell & Laidlaw (1926) and rich cultures were thus regularly made available for the testing of amoebicides *in vitro*. On studying the action of ipecacuanha alkaloids on *E. histolytica* and other entozoic amoeba in this medium, accompanied by a mixed bacterial flora, they then found that emetine and cephaeline, alone of the substances tested, were specifically toxic to *E. histolytica* and concluded that there was a close parallel between the results obtained with *E. histolytica in vitro* and in the treatment of human amoebiasis. Experimental evidence was adduced that emetine acted directly on the amoebae and not on the accompanying bacteria. In a further investigation on the action of emetine on *E. histolytica* in culture, Laidlaw, Dobell & Bishop (1928) showed that the results previously obtained with the biphasic medium were of doubtful value owing to absorption of emetine by the solid phase. A liquid medium was devised consisting of inactivated horse serum in Ringer's fluid (1 : 8) with phosphate buffer to give pH 7·2 and with addition of rice starch. More consistent results in the drug tests were obtained, and it was found that emetine was toxic to *E. histolytica* at dilutions up to 1 in 5 million when incubated with it for 4 days. There were, however, inherent weaknesses in this form of the test, for when its use was extended to a variety of chemical types quite distinct from emetine there was no proof that the amoebicidal effect was not indirect and due to a direct bactericidal action. Subsequent workers have recorded results of amoebicidal tests based on one or other of the tests described above. Some American authors (Anderson & Chuan 1944; Brackett & Bliznick 1947; Rawson & Hitchcock 1947) persist in using the biphasic medium for testing the action of drugs on *E. histolytica* either accompanied by a mixed bacterial flora or a single bacterial organism. It appears to us that their results can have little quantitative significance. Dobell (1947) published an account of an elegant test for direct amoebicidal action *in vitro* in which *E. histolytica* was accompanied by a single species of bacterium, *Bacillus coli*. Methylene blue was added to each tube under test to act as indicator. The tubes were inoculated with a broth culture of *B. coli* and incubated for 24 hr. With good growth of the organism the methylene blue became reduced and, conditions being regarded as suitable for the growth of the amoebae, the tubes were then inoculated. After the desired period of incubation the cultures were examined for growth of amoebae and the concentration of drug at which all had been killed noted, with the aid of subculture in fresh medium if necessary. In the absence of reduction of methylene blue through poor growth of the bacillus due to unsuitable medium or as a result of bactericidal action of the drug, conditions are not suitable for amoebic growth and the test cannot be employed. It was found that at a dilution of 1 in 5 million, emetine killed the amoebae within a period of 4 days.

Of previous attempts to develop an amoebicidal drug superior to emetine, mention must be made of three investigations. Pyman (1937) has given a full account of the extensive researches which he and his collaborators carried out in this field and of which certain aspects only can be touched upon here. Having available a series of *O*-alkylharmols (I) made previously for tests against bird malaria, their amoebicidal activities were examined *in vitro* in the presence of a mixed bacterial flora and were found to show a peak of activity of about 2×10^{-6} at *O-n*-nonylharmol (I, $R = C_9H_{19}$), but this substance suffered from the disadvantage that it was very sparingly soluble in water. The group R was accordingly replaced by a dialkylaminoalkyl radical and maximum activity superior to that found for *O*-nonylharmol was found in the *O*-dibutylaminoundecyl derivative

$$[\text{I, } R = (C_4H_9)_2N . C_{11}H_{22} .].$$
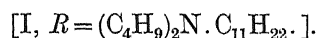
A striking departure from type was then made by omitting the heterocyclic nucleus and terminating the basic side-chain by another basic group; high activity was retained and by extensive exploration of this new type, a maximum of activity was found in 1:10-*bis*(diamylamino)decane (II). The *in vitro* test showed that a concentration of $\frac{1}{3} \times 10^{-6}$ was lethal to *Entamoeba histolytica*. Clinical tests on this substance were disappointing.

(I)   (II)
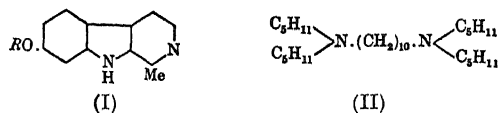
A group of three communications from the Wellcome Laboratories of Tropical Medicine by a team of workers in 1948 may be regarded as a further development of the subject opened up by Pyman. Child & Pyman (1929) had synthesized for test as amoebicides a series of *bis*(6:7-dimethoxytetrahydro*iso*quinolyl)alkanes based on certain fundamental features of the formula of emetine. This line was developed by Goodson and his colleagues (1948*a*) in the second part of the three communications mentioned above. The biological tests by Goodwin, Hoare & Sharp (1948) in the rat, as well as *in vitro* with a single bacterial species, indicated pronounced amoebicidal activity in many of the compounds. In the third communication, Goodson and his colleagues (1948*b*) made many variants of Pyman's *bis*(di-amylamino)decane in which the decane chain was replaced by structures containing benzene nuclei. These compounds by the *in vivo* and *in vitro* tests were on the whole inferior to those described in the previous communication.

The third investigation was carried out in the laboratories of the I.G. Farbenindustrie at Elberfeld, and a connected account has been given by Schönhöfer (1948). Starting from the observation that 2-diethylaminoethylaminodiphenylene oxide (III) had an action on *E. histolytica in vitro*, a great number of variants were made which finally led to the substance 4-*bis*(diethylaminoethyl)amino-3':5'-dimethyldiphenyl ether (IV), known as 'Gavano', which had one-half the toxicity and one-half the activity *in vitro* of emetine. Clinical trials have shown the compound to be slow in action when given orally, and injection was undesirable owing to the pain produced.

In setting up our own organization for screening large numbers of compounds for *in vitro* activity against *E. histolytica*, we have, for ease and rapidity of working, followed the methods introduced by Dobell and his colleagues, in which, in one case, a mixed bacterial flora was present and, in the other, the single organism, *B. coli* accompanied the amoebae. We thereby hoped to obtain sufficient data to enable us to judge whether any correlation existed between the results of the two tests. Jones (1946, 1947) and Goodwin *et al.* (1948) have independently used an experimental infection of *E. histolytica* in the rat for studying the therapeutic action of known and potential amoebicides such as emetine, chiniofon, stovarsol, carbarsone and diodoquin. In our limited experience with this test it appears to suffer from a number of disadvantages. Rats are naturally infected with *E. muris* which complicates diagnosis, and the use of newly weaned rats for intracaecal injections of culture material gives a high mortality frequently associated with bacteriaemia. We have, therefore, not used the rat test as a routine measure.
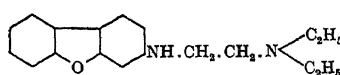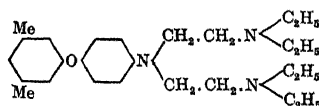
(III)   (IV)

In our own investigations, so far over 400 compounds covering a diversity of organic types have been tested *in vitro*, on *E. histolytica*. When fifty-three compounds had been tested, one, originally supplied during the war by Imperial Chemical Pharmaceuticals Ltd. for test on typhus, namely, 2-*p*-chloroanilino-4-diethylaminoethylaminoquinoline (V), showed activity at a dilution of 10^{-5} using the form of the test devised by Laidlaw *et al.* (1928) in which a mixed bacterial flora is present. This seemed to be a substance which could serve as a basis for chemical development in many directions.
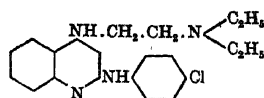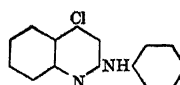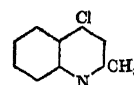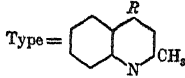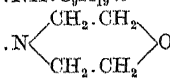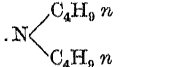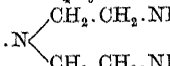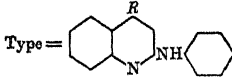
(V)   (VI)   (VII)
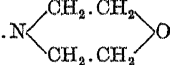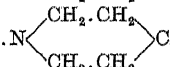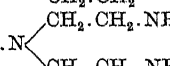
Accordingly, it was decided to start simultaneously from 2-anilino-4-chloroquinoline (VI) and 4-chloroquinaldine (VII) and to build up various types of basic side-chains in place of the reactive chlorine atom in the 4-position of the quinoline nuclei. Unlike Pyman we have retained the quinoline 'carrier' nucleus throughout our compounds. Table 1 shows the results which have been obtained by direct replacement of the halogen atom by primary and secondary amines of various types, the two columns of activities corresponding to amoebicidal tests in presence of (1) a mixed bacterial flora and (2) a single bacterial species. The term 'bactericidal' denotes that a value for amoebicidal activity could not be given because at the highest dilution at which amoebae were absent methylene blue was not reduced.

TABLE 1

Type = (decahydroquinoline with R at position and CH_3 at 2-position)

series no.	$R =$	activity using mixed flora	activity using *B. coli*
A 339	$.NH.CH_3$	10^{-3} to 10^{-4}	not at 10^{-4}
A 368	$.NH.C_3H_7\, n$	10^{-4}	not at 10^{-4}
A 403	$.NH.C_7H_{15}\, n$	10^{-5}	bactericidal
A 353	$.NH.C_8H_{17}\, n$	2×10^{-6} to 10^{-6}	bactericidal
A 404	$.NH.C_9H_{19}\, n$	10^{-5} to 2×10^{-6}	10^{-4} to 10^{-5}
A 252	$.N\langle \frac{CH_2.CH_2}{CH_2.CH_2} \rangle O$	10^{-3}	not at 10^{-4}
A 387	$.N\langle \frac{C_4H_9\, n}{C_4H_9\, n}$	10^{-4} to 10^{-5}	10^{-4} to 10^{-5}
A 382	$.N\langle \frac{CH_2.CH_2.NEt_2}{CH_2.CH_2.NEt_2}$	not at 10^{-3}	not at 10^{-4}
A 236	$.NH.CH_2.CH_2OH$	10^{-3}	not at 10^{-4}
A 237	$.NH.CH_2.CH_2Cl$	10^{-4}	not at 10^{-4}
control	emetine	10^{-6} to 2×10^{-7}	10^{-5} to 2×10^{-7}

Type = (decahydroquinoline with R and NH–cyclohexyl at 2-position)

series no.	$R =$	activity using mixed flora	activity using *B. coli*
A 209	$.NH.C_2H_5$	10^{-5}	bactericidal
A 210	$.NH.C_3H_7\, n$	insoluble	—
A 181	$.N\langle \frac{CH_2.CH_2}{CH_2.CH_2} \rangle O$	not at 10^{-3}	not at 10^{-4}
A 182	$.N\langle \frac{CH_2.CH_2}{CH_2.CH_2} \rangle CH_2$	10^{-3} to 10^{-4}	10^{-4} to 10^{-5}
A 388	$.N\langle \frac{CH_2.CH_2.NEt_2}{CH_2.CH_2.NEt_2}$	10^{-5} to 2×10^{-6}	10^{-4} to 10^{-5}
A 211	$.NH.CH_2.CH_2OH$	10^{-4}	not at 10^{-4}
A 235	$.NH.CH_2.CH_2.Cl$	10^{-5}	bactericidal

As the series of alkylamines is ascended, sparing solubility of the hydrochlorides precludes testing of the compounds beyond the simple ethyl derivative in the 2-anilino series, but the decreased solubility in the quinaldine series causes no difficulty until the octyl derivative is reached. The two compounds A 382 and A 388 are of some interest in that they contain the same side-chain as Gavano (IV), but in these cases the carrier nucleus is different; A 382 is practically inactive, whilst A 388 shows considerable activity. The compounds A 237 and A 235 with reactive halogen atoms are easily accessible and enable one to develop the molecular structure further with retention of water solubility to compounds containing higher alkyl groups than was possible in the parent quinolines. The result of replacing the β-chloro atom in A 237 and A 235 by primary amines is shown in table 2.

TABLE 2

Type = $\begin{array}{c}\text{NH.CH}_2\text{.CH}_2\text{.NH}R\\ \text{[decahydroquinoline ring]}\\ \text{CH}_3\end{array}$
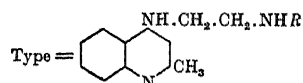
series no.	$R=$	activity using mixed flora	activity using B. coli
A 346	.CH$_3$	10^{-3}	not at 10^{-4}
A 338	.C$_2$H$_5$	10^{-3}	not at 10^{-4}
A 347	.C$_3$H$_7$ n	10^{-4}	not at 10^{-4}
A 227	.C$_3$H$_7$ iso	10^{-4}	not at 10^{-4}
A 240	.C$_4$H$_9$ n	10^{-4}	not at 10^{-4}
A 348	.C$_4$H$_9$ tert.	10^{-4}	not at 10^{-4}
A 365	.C$_5$H$_{11}$ n	10^{-4}	10^{-4}
A 351	.C$_6$H$_{13}$ n	10^{-5}	not at 10^{-4}
A 349	.C$_7$H$_{15}$ n	10^{-6}	bactericidal
A 356	.C$_8$H$_{17}$ n	2×10^{-6} to 10^{-6}	10^{-5} to 2×10^{-6}
A 352	.C$_9$H$_{19}$ n	2×10^{-6} to 10^{-6}	10^{-5} to 2×10^{-6}
A 393	.CH$_2$.CH$_2$.CHMe.CH$_2$.CMe$_3$	10^{-5}	10^{-5} to 2×10^{-6}
A 241	.C$_6$H$_5$	10^{-4}	10^{-4}
A 246	.C$_6$H$_4$Cl p	10^{-4} to 10^{-5}	bactericidal
A 357	.CH$_2$.C$_6$H$_5$	10^{-4} to 10^{-5}	not at 10^{-4}
A 364	.CH$_2$.CH$_2$.C$_6$H$_5$	10^{-5}	10^{-4} to 10^{-5}
A 370	.CH.(CH$_2$)$_4$.CH$_2$	10^{-4}	not at 10^{-4}
control	emetine	10^{-6} to 2×10^{-7}	10^{-5} to 2×10^{-7}

Type = $\begin{array}{c}\text{NH.CH}_2\text{.CH}_2\text{.NH}R\\ \text{[decahydroquinoline ring]}\\ \text{NH—[cyclohexyl]}\end{array}$
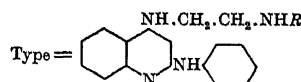
series no.	$R=$	activity using mixed flora	activity using B. coli
A 297	.CH$_3$	10^{-4} to 10^{-5}	10^{-4} to 10^{-5}
A 298	.C$_2$H$_5$	10^{-4} to 10^{-5}	10^{-4} to 10^{-5}
A 309	.C$_3$H$_7$ n	10^{-5}	not at 10^{-4}
A 314	.C$_3$H$_7$ iso	10^{-4} to 10^{-5}	bactericidal
A 310	.C$_4$H$_9$ n	10^{-4} to 10^{-5}	bactericidal
A 313	.C$_4$H$_9$ tert.	10^{-4} to 10^{-5}	not at 10^{-4}
A 325	.C$_5$H$_{11}$ n	10^{-5} to 2×10^{-6}	bactericidal
A 315	.C$_6$H$_{13}$ n	10^{-5} to 2×10^{-6}	bactericidal
A 327	.C$_7$H$_{15}$ n	10^{-6}	10^{-5}
A 251	.C$_6$H$_5$	10^{-4} to 10^{-5}	10^{-4}
A 250	.C$_6$H$_4$Cl p	insoluble	—
A 326	.CH$_2$.C$_6$H$_5$	10^{-5} to 2×10^{-6}	10^{-4} to 10^{-5}
A 369	.CH$_2$.CH$_2$.C$_6$H$_5$	10^{-5}	10^{-4} to 10^{-5}
A 361	.CH.(CH$_2$)$_4$.CH$_2$	10^{-5} to 10^{-6}	bactericidal

This table shows that in the quinaldyl type and the 2-anilinoquinolyl type activity rises as the series is ascended until insolubility of the salts in water precludes further testing. In the quinaldyl series three compounds are lethal to E. histolytica at a dilution of about 1 in 500,000 to 1 in a million. The high activity of the n-octyl member falls off considerably if it is replaced by an eight-carbon radical in the form of the phenylethyl group (A 364); in general branching of the chain and partial or total replacement by cyclic radicals leads to weaker activity. Table 3 demonstrates the effect of replacement of the β-chlorine atom by secondary bases.

TABLE 3

Type = $\text{NH.CH}_2.\text{CH}_2.\text{N}\diagdown^R_R$ (decahydroquinoline with CH_3 at 2-position)
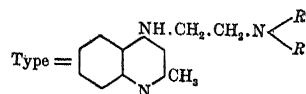
series no.	$R =$	activity using mixed flora	activity using B. coli
A 226	$.\text{CH}_3$	10^{-4}	bactericidal
A 243	$.\text{C}_2\text{H}_5$	10^{-3} to 10^{-4}	not at 10^{-4}
A 225	$.\text{C}_3\text{H}_7\,n$	10^{-4}	bactericidal
A 244	$.\text{C}_3\text{H}_7\,iso$	10^{-3} to 10^{-4}	not at 10^{-4}
A 221	$.\text{C}_4\text{H}_9\,n$	10^{-5}	bactericidal
A 222	$.\text{C}_5\text{H}_{11}\,n$	2×10^{-6}	bactericidal
A 245	$.\text{C}_5\text{H}_{11}\,iso$	2×10^{-6}	bactericidal
A 223	$.\text{C}_6\text{H}_{13}\,n$	10^{-6}	bactericidal
A 378	$.\text{C}_7\text{H}_{15}\,n$	10^{-6}	10^{-5}
A 239	$\text{CH}_2.\text{C}_6\text{H}_5$	10^{-4} to 10^{-5}	10^{-4}
A 238	$\text{CH}_2.\text{CH}_2\text{OH}$	not at 10^{-3}	not at 10^{-4}
A 247	$.\text{CH}_2.\text{CH}_2\text{Cl}$	not at 10^{-3}	10^{-4}
A 379	$.\text{CH}_2.\text{CH}_2.\text{NEt}_2$	not at 10^{-3}	not at 10^{-4}
A 220	$.\text{N}R_2 = $ morpholino	10^{-3}	not at 10^{-4}
A 219	$.\text{N}R_2 = $ piperidino	10^{-4}	not at 10^{-4}
A 366	$.\text{N}R_2 = .\text{N}\diagdown^{\text{CH}_2.\text{CH}_2}_{\text{CH}_2.\text{CH}_2}\diagup\text{N}.\text{C}_6\text{H}_{13}$	10^{-4} to 10^{-5}	10^{-4}
control	emetine	10^{-6} to 2×10^{-7}	10^{-5} to 2×10^{-7}

Type = $\text{NH.CH}_2.\text{CH}_2.\text{N}\diagdown^R_R$ (decahydroquinoline with NH–cyclohexyl at 2-position)
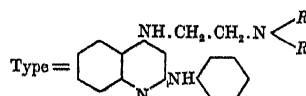
series no.	$R =$	activity using mixed flora	activity using B. coli
A 266	$.\text{CH}_3$	10^{-4} to 10^{-5}	10^{-4} to 10^{-5}
A 234	$.\text{C}_2\text{H}_5$	10^{-5}	bactericidal
A 255	$.\text{C}_3\text{H}_7\,n$	10^{-4} to 10^{-5}	bactericidal
A 233	$.\text{C}_4\text{H}_9\,n$	10^{-5}	bactericidal
A 261	$.\text{C}_5\text{H}_{11}\,n$	insoluble	—
A 256	$.\text{CH}_2.\text{CH}_2\text{OH}$	10^{-4}	not at 10^{-4}
A 354	$.\text{CH}_2.\text{CH}_2\text{Cl}$	10^{-3}	not at 10^{-4}
A 383	$.\text{CH}_2.\text{CH}_2.\text{NEt}_2$	10^{-3}	not at 10^{-4}
A 232	$.\text{N}R_2 = $ morpholino	10^{-4}	10^{-4}
A 218	$.\text{N}R_2 = $ piperidine	10^{-4} to 10^{-5}	bactericidal
A 389	$.\text{CH}_2.\text{CH}_2.\text{N}\diagdown^{\text{CH}_2.\text{CH}_2}_{\text{CH}_2.\text{CH}_2}\diagup\text{O}$	10^{-3}	not at 10^{-4}

Again, a rise of activity in both chemical types is noted, the anticipated smooth gradation of results encountered on ascending the homologous series using the single bacterial organism being confused through the appearance of antibacterial properties in several of the members. Solubility influences again limit the testing in the 2-anilino series sooner than in the quinaldyl series. The characteristic group of 'Gavano', the *bis*(diethylaminoethyl) group, leads to almost complete inactivity (A 379 and A 383).

During the preparation of some of these bases certain *bis*-bases have been isolated as by-products, and one has been prepared from $\gamma\gamma'$-dipiperidyl. No outstanding activity was found in this type as is shown by the results in table 4.

TABLE 4

series no.	R =	activity using mixed flora	activity using B. coli
A 350	.CH$_3$	10^{-3}	not at 10^{-4}
A 242	NR = NN'-$\gamma\gamma'$-dipiperidyl	10^{-4}	bactericidal
A 411	.C$_7$H$_{15}$ n	10^{-4}	not at 10^{-4}
control	emetine	10^{-6} to 2×10^{-7}	10^{-5} to 2×10^{-7}

series no.	R =	activity using mixed flora	activity using B. coli
A 290	.CH$_3$	10^{-4} to 10^{-5}	10^{-4}
A 283	.C$_2$H$_5$	10^{-4} to 10^{-5}	10^{-4}
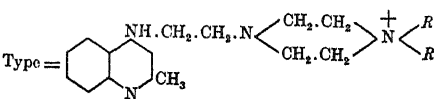
In compounds A 247 and A 354 in table 3 there are two reactive β-chlorine atoms which lend themselves to further substitution. When these compounds were treated with secondary bases like diethylamine and piperidine, the main product was in each case a quaternary salt involving a piperazine ring. In the case of morpholine, however, the quaternary piperazinium salt (VIII) and the ditertiary base (IX) were both isolated:
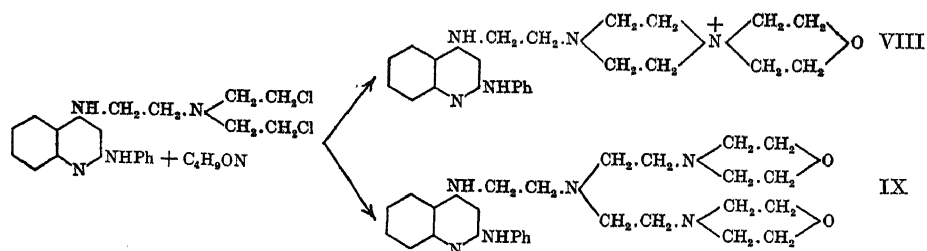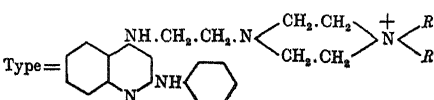
The quaternary compounds showed little activity (table 5). When a primary amine was used instead of a secondary base, a disubstituted piperazine (A 366, table 3) was the normal product.

A point of chemical interest arose in the action of primary amines, e.g. methylamine and benzylamine on 4-β-chloroethylaminoquinaldine and 2-anilino-4-β-chloroethylaminoquinoline respectively. In both cases by-products were formed in which the side-chain in the 4-position was displaced by the reacting amine. Thus methylamine and 4-β-chloroethylaminoquinaldine gave 4-methylaminoquinaldine and 4-β-methylaminoethylaminoquinaldine, whilst benzylamine and 2-anilino-4-β-chloroethylaminoquinoline gave 2-anilino-4-benzylaminoquinoline (X) and 2-

TABLE 5

Type = [decahydroquinoline structure] NH.CH$_2$.CH$_2$.N$\left\langle\begin{array}{c}CH_2.CH_2\\CH_2.CH_2\end{array}\right\rangle\overset{+}{N}\left\langle\begin{array}{c}R\\R\end{array}\right.$ (with CH$_3$ on ring N)

series no.	$R=$	activity using mixed flora	activity using *B. coli*
A 358	.C$_2$H$_5$	not at 10^{-3}	not at 10^{-4}
A 395	$RR=$ .CH$_2$.(CH$_2$)$_3$.CH$_2$.	not at 10^{-3}	not at 10^{-4}
A 248	$RR=$.(CH$_2$)$_2$O(CH$_2$)$_2$.	not at 10^{-3}	not at 10^{-4}
control	emetine	10^{-6} to 2 × 10^{-7}	10^{-5} to 2 × 10^{-7}

Type = [decahydroquinoline structure with N–NH–cyclohexyl] NH.CH$_2$.CH$_2$.N$\left\langle\begin{array}{c}CH_2.CH_2\\CH_2.CH_2\end{array}\right\rangle\overset{+}{N}\left\langle\begin{array}{c}R\\R\end{array}\right.$

series no.	$R=$	activity using mixed flora	activity using *B. coli*
A 363	.C$_2$H$_5$	not at 10^{-3}	not at 10^{-4}
A 362	$RR=$ .CH$_2$.(CH$_2$)$_3$.CH$_2$.	not at 10^{-3}	not at 10^{-4}
A 355	$RR=$.(CH$_2$)$_2$O(CH$_2$)$_2$.	not at 10^{-3}	not at 10^{-4}

anilino-4-β-benzylaminoethylaminoquinoline (XI), the proportion of the former increasing with rise of temperature. In the case of the reaction with benzylamine the side-chain was isolated in the form of benzylaminoethylamine (XII):

[structure] NH.CH$_2$.CH$_2$Cl ... NHPh + C$_6$H$_5$.CH$_2$NH$_2$ $\longrightarrow$ [structure] NH.CH$_2$Ph ... NHPh + [structure] NH.CH$_2$.CH$_2$.NH.CH$_2$Ph ... NHPh + NH$_2$.(CH$_2$)$_2$.NH.CH$_2$Ph

(X)   (XI)   (XII)
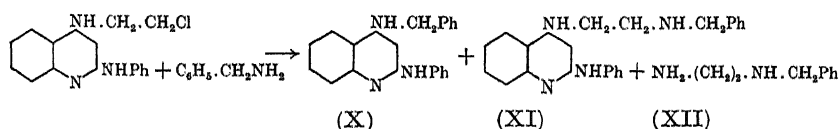
The results of the tests presented in the foregoing tables show that while emetine still retains its pre-eminence as an amoebicidal agent *in vitro*, several of the new synthetic substances possess considerable activity under the same conditions of test and are deemed worthy of further study. It will have been noted that the results obtained by the two test methods show some correlation in so far as a high amoebicidal activity in the specific test with a single bacterial organism is as a rule paralleled by a similar activity in the test with a mixed bacterial flora. It is unfortunate that bactericidal action should be so prominent a feature of the single organism test in one series of compounds (table 3), for although *B. coli* is probably a component of the mixed bacterial flora in the multiple organism test there is no proof that the apparent high amoebicidal action, in the latter test, of the bactericidal compounds is not a real amoebicidal action.

We are left in some doubt as to the relative value of the two types of *in vitro* test used. It is well recognized that where a culture of *E. histolytica* is accompanied by a mixed bacterial flora it may be impossible to tell whether the action of a drug is direct on the amoebae or indirect through its action on one or more of the essential bacteria present as may be indicated by a lack of turbidity in the culture. This test is easier to carry out than the other, since amoebae grow more abundantly

and a more satisfactory end-point can be obtained. Conditions in the culture, moreover, approach more closely those in the human intestine from which the material was originally isolated. The second type of test we have used has fewer uncontrolled factors and a specific action on the amoebae can be detected. It suffers from the drawback, however, that a substance bactericidal to *B. coli* cannot be used in the test, whereas such a property might be of the utmost value in the treatment of amoebiasis (Hargreaves 1945). So far as we are aware, no new amoebicides of value have been discovered hitherto by *in vitro* or animal tests, and it may be premature to give a categorial statement on the value of the methods.

EXPERIMENTAL

Biological

The strain of *E. histolytica* used throughout the first series of tests was obtained from a human case of amoebiasis and, accompanied by a mixed flora, has grown well in the liquid medium described by Laidlaw *et al.* (1928). In preliminary trials each drug, dissolved in the same medium, was made up in pairs of tubes at concentrations of 10^{-3}, 10^{-4} and 10^{-5} in 5 ml. amounts. If insoluble in the medium the solution of 20 mg. of substance was first effected in water, generally 1 ml., and then mixed with 19 ml. of medium, when clear solutions or fine suspensions resulted. Seven drugs were as a rule used in a single experiment along with six control tubes, making a total of forty-eight tubes, and in addition emetine was frequently tested at the same time. When necessary the range of drug dilutions was extended so that a clear-cut result was obtained. After addition of a loopful of rice starch, the tubes containing drugs and the controls were incubated overnight to check for sterility. On the following day, 0·1 ml. of a suspension of amoebae from stock cultures, containing at least one parasite per field in a fresh film preparation, was added to each, and they were then incubated at 37° C for 3 days. At the end of this period control cultures were first examined, and if growth was satisfactory each of two observers examined microscopically fresh preparations from one tube from each pair of the different drug dilutions using a 6 times eyepiece and ⅓ in. objective; the results were then compared. In cases of doubt subcultures were made to determine if live amoebae were present. The lethal concentration of a drug was recorded as the highest dilution at which all amoebae were killed. In the case of some drugs a bactericidal action was evidenced by absence of turbidity in the cultures and note was made of this fact.

In the second series of tests a culture of *E. histolytica* was used in which the only other organism present was a non-motile form of *B. coli*. For this culture we were indebted to the late Clifford Dobell, F.R.S. In preliminary trials with this strain we frequently found, in agreement with the original author, that cultures grew badly or not at all in the liquid medium described earlier. We therefore enriched it by the addition of 20 % Wright's broth (1933) or liver extract and satisfactory growth was then consistently obtained. The lethal action of emetine was then compared in the three different media. In the case of the original medium and that with broth the results were for practical purposes the same, but in the liver medium

the lethal action of emetine was reduced and we have not used it in our tests. Instead of adding 1 drop of methylene-blue solution to each tube, it was found more convenient to add 10 ml. of a 0·1 % solution of the substance to 1 l. of the medium before distribution. The *B. coli* used to inoculate the culture tubes was grown on a serum-agar slope and suspended in 2 % glucose broth, so that 30×10^8 organisms were present per ml. as indicated by comparison with an opacity standard. One drop of the suspension was added to each culture tube, and in practice the resultant growth proved satisfactory for reduction of methylene blue without at the same time giving rise to overgrowth during the experimental period. The test was otherwise carried out as described by Dobell (1947) and emetine was always included as a standard. The lowest dilution of any drug used in this case was 10^{-4}.

Chemical

2:4-Dichloroquinoline. This substance was prepared following the method of Brooker & Smith (1937) (see, however, Buchmann & Hamilton 1942). The product distilled at 158° C at 15 mm. · It was then most conveniently crystallized from spirit and gave an 85 % yield, m.p. 60° C. (Brooker & Smith recommend methyl alcohol for its crystallization and record m.p. 66 to 67° C in agreement with Baeyer & Bloem (1882) who gave m.p. 67° C.)

The residue from the distillation gave a fraction b.p. 180 to 200° C at 20 mm. which crystallized and on solution in ethyl alcohol gave a trichloroquinoline, in long asbestos-like needles, m.p. 107° C. (Found: C, 46·4; H, 1·8; N, 6·3. Calc. for $C_9H_4NCl_3$, C, 46·5; H, 1·7; N, 6·0 %.) This may be identical with the trichloroquinoline of Rügheimer (1884, 1885), to which the constitution 2:3:4-trichloroquinoline is ascribed without rigid proof for the position 3 for one of the chlorine atoms.

A. Compounds derived from 2-anilino-4-chloroquinoline

2-Anilino-4-chloroquinoline. This substance was prepared by the method of Curd, Raison & Rose (1947), by their procedure *a* in 59 % yield, m.p. 164° C. The recrystallization mother liquors contain other substances, among which 2-hydroxy-4-chloroquinoline and 2:4-dianilinoquinoline hydrochloride (Dziewonski & Dymek 1937) have been identified.

2-Anilino-4-ethylaminoquinoline hydrochloride. 2-Anilino-4-chloroquinoline (5·1 g.) and alcoholic ethylamine (8 ml.; 33 %) were heated in a sealed tube at 200° C for 6 hr. Alcohol and free ethylamine were pumped off, the residue made alkaline and extracted with ether. On removing the ether the partly crystalline residue was converted into the *hydrochloride* which on solution in alcohol (20 ml.) crystallized in short rods (3·3 g.), m.p. 253 to 254° C. (Found: C, 68·0; H, 6·0; N, 13·9. $C_{17}H_{17}N_3$,HCl requires C, 68·1; H, 6·0; N, 14·0 %.)

2-Anilino-4-n-propylaminoquinoline hydrochloride. 2-Anilino-4-chloroquinoline (5·1 g.) and *n*-propylamine (3·5 g.) were heated together as in the preceding experiment. The required *hydrochloride* (2·2 g.) crystallized from alcohol in clusters of prisms, m.p. 246° C. (Found: C, 69·1; H, 6·5; N, 13·4. $C_{18}H_{19}N_3$, HCl requires C, 68·9; H, 6·4; N, 13·4 %.)

2-Anilino-4-n-butylaminoquinoline hydrochloride. 2-Anilino-4-chloroquinoline (2·54 g.) and n-butylamine (2·13 g.) gave in the usual way the *hydrochloride* which crystallized from alcohol in tablets, 2·3 g., m.p. 219° C. (Found: C, 70·0; H, 7·0; N, 13·2. $C_{19}H_{21}N_3$, HCl requires C, 69·6; H, 6·8; N, 12·8 %.) This and the preceding salt were not sufficiently soluble in water for biological tests.

2-Anilino-4-bis(diethylaminoethyl)aminoquinoline trihydrobromide. 2-Anilino-4-chloroquinoline (2·54 g.) and *bis*(diethylaminoethyl)amine (5·4 g.) were heated at 205° C for 3 hr. N-Sodium hydroxide (20 ml.) was added and the mixture warmed on the water-bath, then cooled in ice and decanted from the gum; this process was repeated three times more to remove excess of water-soluble amine and the residue dissolved in alcohol and heated with N-hydrobromic acid (35 ml.), when the solution reacted acid to Congo paper. The solution was evaporated dry, the residue dissolved in absolute alcohol and on keeping the *trihydrobromide* (2 g.) crystallized in clusters of tiny needles. It was recrystallized from absolute alcohol (32 ml.) for analysis, yield 1·65 g., m.p. 259 to 260° C. (Found: C, 47·3; H, 6·4; N, 10·5. $C_{27}H_{39}N_5$, 3HBr requires C, 47·9; H, 6·3; N, 10·4 %.)

2-Anilino-4-morpholinoquinoline hydrochloride. Morpholine (10 ml.) and 2-anilino-4-chloroquinoline were heated in an oil-bath at 130° C for 6 hr. The free morpholine was pumped off with warming and the residue treated with 2N-sodium hydroxide solution and ether. The sparingly soluble solid (1·4 g.), was collected and crystallized from alcohol (60 ml.) from which 2-*anilino-4-morpholinoquinoline* separated in long silky needles, m.p. 202° C. (Found: C, 74·1; H, 6·1; N, 13·8. $C_{19}H_{19}ON_3$ requires C, 74·7; H, 6·2; N, 13·8 %.) The ethereal solution gave a further quantity, 0·5 g., of the base. The *hydrochloride* separated from water in pointed prisms. (Found: N, 12·4; $C_{19}H_{19}ON_3$, HCl requires N, 12·3 %.)

2-Anilino-4-piperidinoquinoline hydrochloride. 2-Anilino-4-chloroquinoline (5·08 g.) and piperidine (8·5 g.) were heated at 170° C in a sealed tube for 6 hr. The free piperidine was pumped off, the residue made alkaline and extracted with ether. The residue (8·1 g.) from the ether after removing piperidine was dissolved in spirit (50 ml.) and 2-*anilino-4-piperidinoquinoline* (4·7 g.) separated in needles, m.p. 154° C. (Found: N, 13·9. $C_{20}H_{21}N_3$ requires N, 13·9 %.) The *hydrochloride* separated from water in prisms, m.p. 244° C. (Found: C, 70·6; H, 6·4; N, 12·4. $C_{20}H_{21}N_3$, HCl requires C, 70·7; H, 6·2; N, 12·4 %.)

2-Anilino-4-benzylaminoquinoline hydrochloride. 2-Anilino-4-chloroquinoline (2·54 g.) and benzylamine (10·7 g., 10 mol.) were heated to 170 to 175° C for 4 hr. Aqueous alkali was added and the base collected in chloroform; the benzylamine was completely removed by distillation at 15 mm. The residue was dissolved in spirit (5 ml.) and treated with N-hydrochloric acid (15 ml.). The crystalline *hydrochloride* separated in needles, yield 3·18 g. For analysis it was crystallized from absolute alcohol (25 ml.) and had m.p. 244° C. (Found: C, 72·9; H, 5·9. $C_{22}H_{19}N_3$, HCl requires C, 73·0; H, 5·6 %.)

2-Anilino-4-phenoxyquinoline. 2-Anilino-4-chloroquinoline (2·5 g.), phenol (5 g.) and dibutylamine (4·3 ml.) were heated in a sealed tube at 210° C for 6 hr. The free dibutylamine was pumped off, the residue made alkaline and repeatedly extracted with chloroform. On removing the chloroform the residue crystallized from alcohol

in long prisms (1·6 g.), m.p. 157 to 158° C. (Found: C, 80·6; H, 5·0; N, 9·4. $C_{21}H_{16}ON_2$ requires C, 80·7; H, 5·2; N, 9·0 %.)

2-*Anilino-4-hydroxyethylaminoquinoline*. 2-Anilino-4-chloroquinoline (25·4 g., 1 mol.) and ethanolamine (12·2 ml., 2 mol.) were mixed and heated in an oil-bath to 180 to 185° C (internal temp.) when a vigorous reaction took place. The flask was temporarily raised from the bath until the reaction had subsided. The contents were then heated at 190° C for 2 hr., dissolved in a small volume of alcohol and poured into a well-stirred solution of weak alkali. The solid was collected, ground up with 2N sodium hydroxide solution until free from chloride. The yield was 26·8 g., m.p. 142° C. It was crystallized from spirit (40 ml.) and separated in flattened prisms, m.p. 149° C. (Found: C, 72·6; H, 6·1; N, 15·4. $C_{17}H_{17}ON_3$ requires C, 73·0; H, 6·1; N, 15·1 %.)

The combined mother liquors from many similar batches gave 2:4-*diethanolaminoquinoline*, m.p. 192° C, as clusters of needles from spirit. (Found: C, 63·1; H, 6·8; N, 17·0. $C_{13}H_{17}O_2N_3$ requires C, 63·1; H, 6·9; N, 17·0 %.)

2-*Anilino-4-hydroxyethylaminoquinoline hydrochloride* crystallized from water in tablets, m.p. 234 to 235° C when dried at 100° C. (Found: C, 61·8; H, 6·0; N, 13·0; H_2O, 4·8. $C_{17}H_{17}ON_3$, HCl, H_2O requires C, 61·2; H, 6·0; N, 12·6; H_2O, 5·4 %.)

2-*Anilino-4-chloroethylaminoquinoline hydrochloride*. 2-Anilino-4-hydroxyethylaminoquinoline (8·4 g.) in benzene (20 ml.) was treated with thionyl chloride (2·19 ml.) and heated on a water-bath for 4 hr. The solid was collected, yield 10·6 g., m.p. 242 to 248° C., and crystallized from 50 % alcohol (200 ml.) and separated in tufts of long woolly needles, m.p. 252° C. (Found: C, 60·9; H, 5·1; N, 12·7. $C_{17}H_{16}N_3Cl$, HCl requires C, 61·0; H, 5·1; N, 12·6 %.)

2-*Anilino-4-methylaminoethylaminoquinoline dihydrochloride*. The preceding chloroethyl-hydrochloride (3·34 g., M/100) and alcoholic methylamine (12 ml. 33 %, M/10) were heated in a sealed tube at 150° C for 4 hr. The base was converted into the *dihydrochloride* which was crystallized twice from absolute alcohol by adding ethyl acetate. It separated in fine needles, yield 2·1 g., m.p. 103° C. If dried at 100° C its m.p. is indefinite. (Found: H_2O, 8·4. $C_{18}H_{20}N_4$, 2HCl, $2H_2O$ requires H_2O, 9·0 %. On dried solid, C, 58·6; H, 6·0; N, 15·0. $C_{18}H_{20}N_4$, 2HCl requires C, 59·2; H, 6·1; N, 15·3 %.)

2-*Anilino-4-ethylaminoethylaminoquinoline dihydrochloride*. This *salt* was prepared in a similar way to the preceding. From alcohol-ethyl acetate it separated in woolly needles, yield 2·2 g., m.p. below 100° C. When dried *in vacuo* it melts between 150 and 155° C with considerable shrinking at a lower temperature. (Found: C, 55·2; H, 6·6; N, 13·2; H_2O, 8·3. $C_{19}H_{22}N_4$, 2HCl, $2H_2O$ requires C, 54·9; H, 6·8; N, 13·5; H_2O, 8·7 %.)

2-*Anilino-4-n-propylaminoethylaminoquinoline dihydrochloride*. Prepared in a similar way to the two preceding salts, this *dihydrochloride* crystallized from absolute alcohol in tiny needles, yield 2·6 g., m.p. 263 to 264° C. (Found: C, 61·5; H, 6·4; N, 14·3. $C_{20}H_{24}N_4$, 2HCl requires C, 61·1; H, 6·7; N, 14·3 %.) The nitrate, needles, is much less soluble.

2-*Anilino-4-isopropylaminoethylaminoquinoline dihydrochloride*. Prepared in a similar way to the three preceding salts, this *dihydrochloride* crystallized from

absolute alcohol in short needles, yield 3·3 g., m.p. 263 to 265° C. (Found: C, 61·4; H, 6·3; N, 14·1. $C_{20}H_{24}N_4$, 2HCl requires C, 61·1; H, 6·2; N, 14·2 %.) This substance also crystallizes as a *dihydrate* in large well formed prisms, m.p. 236 to 237° C. (Found: C, 56·1; H, 6·7; N, 13·4; H_2O, 8·7. $C_{20}H_{24}N_4$, 2HCl, $2H_2O$ requires C, 55·9; H, 7·0; N, 13·0; H_2O, 8·4 %.)

2-Anilino-4-butylaminoethylaminoquinoline dihydrobromide. Prepared in a similar way to the above, this *dihydrobromide*, 2·0 g., m.p. 242° C, crystallized from absolute alcohol in needles. (Found: C, 51·2; H, 5·4; N, 11·4. $C_{21}H_{26}N_4$, 2HBr requires C, 50·8; H, 5·7; N, 11·3 %.)

2-Anilino-4-tert-butylaminoethylaminoquinoline dihydrobromide. This salt was obtained in a similar way; the yield of small needles from spirit was 2·3 g., m.p. 280° C. (Found: C, 50·6; H, 5·2; N, 10·8. $C_{21}H_{26}N_4$, 2HBr requires C, 50·8; H, 5·7; N, 11·3 %.)

2-Anilino-4-n-amylaminoethylaminoquinoline dihydrochloride. This salt, prepared in a similar way to the above, crystallized from absolute alcohol in woolly needles (2·53 g.) which are an unstable form; in contact with the solution they soon change into microscopic tablets or prisms, m.p. 204 to 205° C. (Found: C, 60·1; H, 7·2; N, 12·7; H_2O, 4·3. $C_{22}H_{28}N_4$, 2HCl, H_2O requires C, 60·1; H, 7·3; N, 12·8; H_2O, 4·1 %. On anhydrous salt, Cl = 16·2. $C_{22}H_{28}N_4$, 2HCl requires Cl, 16·8 %.)

2-Anilino-4-n-hexylaminoethylaminoquinoline dihydrobromide. This salt, prepared similarly, crystallizes from absolute alcohol on addition of ethyl acetate in microscopic needles (3·0 g.) which are hygroscopic. (Found: Br, 30·5. $C_{23}H_{30}N_4$, 2HBr requires Br, 30·5 %.)

2-Anilino-4-n-heptylaminoethylaminoquinoline dihydrochloride. This salt crystallizes from alcohol-ethyl acetate in woolly needles (2·9 g.). It is difficultly soluble in water. (Found: N, 12·2; H_2O, 2·3. $C_{24}H_{32}N_4$, 2HCl, 0·5 H_2O requires N, 12·2; H_2O, 2·0 %. On the salt dried at 100° C, Cl = 16·1. $C_{24}H_{32}N_4$, 2HCl requires Cl, 15·8 %.)

2-Anilino-4-β-anilinoethylaminoquinoline. 2-Anilino-4-chloroethylaminoquinoline hydrochloride (3·34 g.) and aniline (5 ml.) were heated in an oil-bath kept at 180° C for 1·5 hr. The solution was made alkaline and steam distilled to free from aniline. On cooling the aqueous residue the gum readily crystallized, yield 4·1 g. It was warmed with N-hydrochloric acid (30 ml.) and the *dihydrochloride* collected (4·0 g.). It was dissolved in *iso*propyl alcohol (80 ml.) and separated in narrow plates, m.p. 215° C (efferv.). (Found: C, 63·3; H, 6·1; N, 13·3; Cl, 15·9; H_2O, 2·2. $C_{23}H_{22}N_4$, 2HCl, 0·5 H_2O requires C, 63·3; H, 5·8; N, 12·9; Cl, 16·3; H_2O, 2·1 %.)

2-Anilino-4-β-p-chloroanilinoethylaminoquinoline dihydrochloride. The chloro-base hydrochloride (3·34 g.) and *p*-chloroaniline were heated at 190° C for 3 hr. The crystalline base, m.p. 104° C, long needles from methyl alcohol, was converted into the *dihydrochloride* which crystallized from alcohol in rectangular flat prisms, m.p. 247° C. (Found: C, 64·0; H, 5·3; N, 12·7; H_2O, 1·9. $C_{23}H_{21}N_4Cl$, 2HCl, 0·5 H_2O requires C, 63·6; H, 5·3; N, 12·9; H_2O, 2·1 %.)

2-Anilino-4-benzylaminoethylaminoquinoline dihydrobromide. The parent chloroethyl-hydrochloride (6·68 g., M/50) and benzylamine (7·0 g.) were heated at 130° C in an oil-bath for 4 hr. The product was worked up as the base and the excess of

benzylamine removed. The residual base was converted into the *dihydrobromide* and crystallized on prolonged keeping at 0° C from absolute alcohol-ethyl acetate in tiny needles (6·2 g.). Once it had been isolated in this way it crystallized from absolute alcohol in needles or prisms which melt below 120° C. If dried at 100° C it melts about 134° C but only forms a meniscus at much higher temperatures. (Found: C, 52·3; H, 5·5; N, 10·4; H_2O, 4·4. $C_{24}H_{24}N_4$, 2HBr, 1·5 H_2O requires C, 52·6; H, 5·3; N, 10·2; H_2O, 4·9 %. On salt, dried at 100° C, Br, 30·1. $C_{24}H_{24}N_4$, 2HBr requires Br, 30·7 %.)

When this reaction is carried out at 150° C a mixture of bases is obtained through some replacement of the side-chain by the benzyl radical. The side-chain appears as benzylethylenediamine which was isolated as its dihydrochloride, pearly leaflets from ethyl alcohol, m.p. 260 to 261° C (Bleier (1899), gives m.p. 253° C). (Found: C, 48·2; H, 6·9; N, 12·5. Calc. for $C_9H_{14}N_2$, 2HCl, C, 48·4; H, 7·2; N, 12·6 %.) The 2-anilino-4-benzylaminoquinoline hydrochloride and 2-anilino-4-benzylamino-ethylaminoquinoline dihydrochloride also present were inseparable in most experiments but occasionally both could be isolated and proved to be identical with the above described salts.

2-Anilino-4-cyclohexylaminoethylaminoquinoline dihydrobromide. 2-Anilino-4-chloroethylaminoquinoline hydrochloride (3·34 g.) and *cyclo*hexylamine (9·9 g., 10 mol.) were heated at 120° C for 5 hr. When cold 2N-sodium hydroxide was added and the bases extracted with chloroform. After removing the chloroform, excess of *cyclo*hexylamine was removed by heating in a vacuum in the boiling water-bath and the residue treated with slight excess of 3N-hydrobromic acid in presence of alcohol. The required *dihydrobromide* crystallized in clusters of small prisms, yield 4·4 g. It was twice crystallized from methyl alcohol and separated as a monohydrate, m.p. indefinite between 160° and 170° C. If however it is dissolved in absolute ethyl alcohol it rapidly crystallizes out as the anhydrous form which is then very sparingly soluble and melts at 257° C. (Found: C, 50·7; H, 5·9; N, 10·5; H_2O, 3·7. $C_{23}H_{28}N_4$, 2HBr, H_2O requires C, 51·1; H, 6·0; N, 10·4; H_2O, 3·3 %.)

2-Anilino-4-β-phenylethylaminoethylaminoquinoline dihydrobromide. The usual chloroethylamine (3·34 g.) and β-phenylethylamine (6·05 g., 5 mol.) were heated at 120 to 125° C for 3 hr. The chloroform extract of the alkalinized solution was freed from chloroform and phenylethylamine by distillation and the residual base treated with hydrobromic acid in excess and evaporated to dryness. The residue dissolved in absolute alcohol (10 ml.) and gave the required *dihydrobromide* (3·85 g.) as clusters of prisms. After another crystallization it had m.p. 195 to 196° C. (Found: C, 54·8; H, 5·0; N, 10·0. $C_{25}H_{26}N_4$, 2HBr requires C, 55·1; H, 5·2; N, 10·3 %.)

2-Anilino-4-dimethylaminoethylaminoquinoline. The chloro-compound (3·34 g.) and aqueous dimethylamine (10 ml., 33 %) were heated at 140° C in a sealed tube for several hours. The *base* was isolated and crystallized from 85 % alcohol in columns, m.p. 86 to 88° C but when dried *in vacuo* at 122° C. (Found: C, 69·8; H, 7·3; N, 17·5; H_2O, 5·9. $C_{19}H_{22}N_4$, H_2O requires C, 70·3; H, 7·5; N, 17·3; H_2O, 5·6 %.) It was converted into the *dihydrobromide* which crystallized from alcohol

in needles, yield 4·15 g., m.p. 225° C. (Found: C, 46·1; H, 5·4; N, 11·8; H_2O, 5·9. $C_{19}H_{22}N_4$, 2HBr, 1·5 H_2O requires C, 46·1; H, 5·1; N, 11·3; H_2O, 5·5. On dried solid, Br, 34·0. $C_{19}H_{22}N_4$, 2HBr requires Br, 34·1 %.) The *dihydrochloride* separated from ethyl alcohol in clusters of prisms, m.p. 219 to 220° C. (Found: N, 13·7; H_2O, 5·5. $C_{19}H_{22}N_4$, 2HCl, H_2O requires N, 14·1; H_2O, 4·5 %. On dried solid, Cl, 18·5. $C_{19}H_{22}N_4$, 2HCl requires Cl, 18·7 %.)

2-Anilino-4-diethylaminoethylaminoquinoline dihydrochloride. This *salt* was prepared in a similar way using aqueous diethylamine. It crystallized from ethyl alcohol on addition of three volumes of ethyl acetate in small needles, m.p. 256° C. (Found: C, 59·5; H, 7·2; N, 13·4; H_2O, 4·6. $C_{21}H_{26}N_4$, 2HCl, H_2O requires C, 59·3; H, 7·1; N, 13·2; H_2O, 4·2 %.)

2-Anilino-4-β-n-dipropylaminoethylaminoquinoline dihydrobromide. The chloro-base hydrochloride (3·34 g.) and *n*-dipropylamine (6·7 ml.) were heated in a sealed tube at 180° C for a few hours. The base was isolated in the usual way and converted into the *dihydrobromide* which crystallized from absolute alcohol (2·5 ml.) on addition of dry ethyl acetate (6 ml.) in clusters of plates, yield 2·38 g., m.p. indefinitely between 135 and 155° C. (Found: C, 50·4; H, 6·3; N, 10·4; Br, 28·9; H_2O, 4·3. $C_{23}H_{30}N_4$, 2HBr, H_2O requires C, 50·9; H, 6·3; N, 10·3; Br, 29·5; H_2O, 3·3 %.)

2-Anilino-4-n-dibutylaminoethylaminoquinoline dihydrochloride. The chloro-compound (3·34 g.) and dibutylamine (6·5 ml.) were heated in a sealed tube at 140° C, and the required base isolated in the usual way. It was converted into the *dihydrochloride* which crystallized from alcohol on addition of ethyl acetate in prisms, yield 3·5 g., m.p. 118 to 120° C. (Found: C, 61·5; H, 8·0; N, 11·6; H_2O, 8·1. $C_{25}H_{34}N_4$, 2HCl, 1·5 H_2O requires C, 61·2; H, 8·0; N, 11·4; H_2O, 5·5 %.)

2-Anilino-4-β-di-n-amylaminoethylaminoquinoline hydrobromide. The β-chloro-base hydrochloride (3·34 g., M/100) and *n*-diamylamine (7·9 g., M/20) were heated in a bath at 185 to 190° C for several hours. After making alkaline and steam distilling, the residual base was extracted with chloroform, the solvent removed and the residue converted into the *monohydrobromide* which crystallized from absolute alcohol in clusters of needles, yield 2·0 g., m.p. 137 to 138° C. (Found: C, 65·0; H, 8·6; N, 10·9; Br, 16·5. $C_{27}H_{38}N_4$, HBr requires C, 64·9; H, 7·9; N, 11·2; Br, 16·0 %.) This salt does not dissolve completely in water, probably owing to partial hydrolysis.

2-Anilino-4-bis(diethylaminoethyl)aminoethylaminoquinoline 3·5 hydrobromide. The β-chlorobase hydrochloride (3·34 g.) and bis(diethylaminoethyl)amine (10·75 g.) were heated at 120° C for 3 hr. When cool, 2N sodium hydroxide (15 ml.) and water (10 ml.) were added, the mixture warmed on the water-bath, then cooled in ice to render the water-insoluble material viscous and the aqueous layer decanted. Four more extractions each with water (10 ml.) were made similarly. The residual gum was dissolved in alcohol, N-hydrobromic acid (45 ml.) added and the solution evaporated to dryness. The crystalline residue in boiling absolute alcohol (75 ml.) only again crystallized on the addition of water (2·5 ml.), yield 5·15 g. It was recrystallized from spirit (10 ml.) and separated in needles, m.p. 135 to 136° C. (Found: C, 45·8; H, 6·7; N, 11·0. $C_{29}H_{44}N_6$, 3·5 HBr requires C, 45·8; H, 6·3;

N, 11·0 %.) This salt gives a neutral solution in water. When crystallized in presence of excess of hydrobromic acid, it separates as a *tetrahydrobromide*, m.p. 231 to 233° C, which crystallizes in microscopic needles. (Found: C, 43·9; H, 6·1; N, 10·2. $C_{29}H_{44}N_6$, 4HBr requires C, 43·5; H, 6·1; N, 10·5 %.) The aqueous solution of this salt is acid in reaction.

2-Anilino-4-diethanolaminoethylaminoquinoline dihydrochloride. The components in the usual proportions were heated at 180° C for several hours. The product was warmed with water, then cooled and the aqueous liquor decanted, this process being repeated three times. The oily residue was dissolved in excess of hydrochloric acid and the solution evaporated to dryness. The *dihydrochloride* was dissolved in absolute alcohol (10 ml.) and separated in needles which on one further crystallization had m.p. 219 to 220° C, yield 2·84 g. (Found: on air-dried solid, C, 55·8; H, 6·7; N, 12·4; H_2O, 2·8. $C_{21}H_{26}O_2N_4$, 2HCl, $\frac{3}{4}H_2O$ requires C, 55·7; H, 6·6; N, 12·4; H_2O, 3·0. On solid dried at 100° C, Cl, 16·1. $C_{21}H_{26}O_2N_4$, 2HCl requires Cl, 16·1 %.)

2-Anilino-4-β-di(chloroethyl)aminoethylaminoquinoline dihydrochloride. 2-Anilino-4-diethanolaminoethylaminoquinoline dihydrochloride (11·5 g., anhydrous salt) was boiled in suspension in benzene with thionyl chloride (7·5 ml., 4 mol.) for 2 hr. The solid was collected (12·3 g.) and crystallized from spirit (40 ml.) from which it separated in colourless needles, yield 10·2 g., m.p. 192 to 193° C. (Found on air-dry material, loss at 100° C, 3·3. $C_{21}H_{24}N_4Cl_2$, 2HCl, H_2O requires H_2O, 3·6 %. On anhydrous salt, C, 52·9; H, 5·9. $C_{21}H_{24}N_4Cl_2$, 2HCl requires C, 52·9; H, 5·5 %.)

2-Anilino-4-piperidinoethylaminoquinoline. The chloro-ethyl hydrochloride (3·34 g.) and piperidine (5 ml.) were heated at 130° C for 3 hr. The product was treated with 2N-sodium hydroxide solution and the solid (4·2 g.) collected, dried and crystallized from alcohol when the required *base* separated in needles, m.p. 139 to 140° C. (Found: C, 76·3; H, 7·3; N, 16·6. $C_{22}H_{26}N_4$ requires C, 76·3; H, 7·6; N, 16·2 %.) The *dihydrochloride* separated from N-hydrochloric acid in slender prisms. (Found: C, 61·6; H, 6·7; N, 13·2; H_2O, 1·2. $C_{22}H_{26}N_4$, 2HCl, 0·5 H_2O requires C, 61·6; H, 6·8; N, 13·1; H_2O, 2·1 %.)

2-Anilino-4-morpholinoethylaminoquinoline. Prepared in a similar way to the preceding, this *base* crystallized from alcohol in delicate rhomb-shaped leaflets, m.p. 168° C. (Found: C, 72·2; H, 7·2; N, 16·2. $C_{21}H_{24}ON_4$ requires C, 72·4; H, 7·0; N, 16·1 %.) The *dihydrochloride* separated from alcohol in woolly needles, m.p. 138 to 140° C. (Found: C, 54·7; H, 6·5; N, 12·2; H_2O, 8·6. $C_{21}H_{24}ON_4$, 2HCl, 2H_2O requires C, 55·1; H, 6·6; N, 12·3; H_2O, 7·9 %.)

Bis-(2-anilino-4-quinolylaminoethyl)methylamine. The parent chloroethyl-hydrochloride (3·34 g.) and aqueous methylamine (10 ml., 33 %) were heated at 130° C in a sealed tube for 5 hr. The basic products were recovered in chloroform and on solution in methyl alcohol gave the bis-*base* (1·3 g.) m.p. 194° C crystallizing in prisms. This base, unlike the secondary base, is sparingly soluble in methyl alcohol. (Found: C, 75·2; H, 6·3; N, 17·6. $C_{35}H_{35}N_7$ requires C, 75·9; H, 6·4; N, 17·7 %.)

Attempts to increase the proportion of this *bis*-base by reducing the proportion of methylamine used and working in alcoholic media were no more successful. Examination of the mother liquors in each case showed the presence of the secondary base by isolation of the previously described dihydrochloride.

The *trihydrochloride* of the *bis*-base crystallizes from spirit in small needles, m.p. indefinite between 200° and 210° C. (Found: on macro-analysis, loss at 100° C, 7·6%. $C_{35}H_{35}N_7$, 3HCl, $3H_2O$ requires H_2O, 7·5%. On micro-analysis, loss at 100° C, 6·4. $C_{35}H_{35}N_7$, 3HCl, $3H_2O$ losing 2·5 H_2O requires loss 6·3%. On this dried sample, found: C, 62·2; H, 5·9; N, 14·3. $C_{35}H_{35}N_7$, 3HCl, 0·5 H_2O requires C, 62·5; H, 5·9; N, 14·6%.)

Bis-(2-*anilino-4-quinolylaminoethyl*)*ethylamine*. The parent chloroethyl-hydrochloride (3·34 g.) and aqueous ethylamine (10 ml., 33%) were heated at 130° C for 5 hr. The basic products extracted with chloroform readily crystallized on removing the solvent and on crystallization from boiling methyl alcohol (80 ml.) separated in needles which soon became opaque in appearance (1·2 g.), m.p. 174 to 175° C. (Found: C, 76·1; H, 6·6; N, 17·3. $C_{36}H_{37}N_7$ requires C, 76·2; H, 6·6; N, 17·3%.) The *trihydrochloride* was soluble in 20 parts of boiling spirit and crystallized in prisms or needles, m.p. 262° C (efferv.). (Found: C, 59·9; H, 5·9; N, 13·9; H_2O, 5·0. $C_{36}H_{37}N_7$, 3HCl, $2H_2O$ requires C, 60·6; H, 6·2; N, 13·8; H_2O, 5·0. On the solid dried at 100° C, Cl, 15·5. $C_{36}H_{37}N_7$, 3HCl requires Cl, 15·7%.)

N-(2-*Anilino-4-quinolylaminoethyl*)-*N′N′*-*diethylpiperazinium bromide dihydrobromide*. The dichloro-base dihydrochloride (2·38 g. of anhydrous salt) was heated in a sealed tube with diethylamine (2·6 ml., 5 mol.) and absolute alcohol (5 ml.) at 115° C for 5 hr. The alcohol was removed, 2N-sodium hydroxide and chloroform added and the middle oily layer collected as in the two following experiments. The oily layer was evaporated to dryness with 3N-hydrobromic acid and the crystalline residue recrystallized from 20 parts of boiling spirit from which it separated in needles, yield 2·1 g., m.p. 246 to 247° C. (Found: C, 44·1; H, 5·6; N, 10·4; Br, 35·7; H_2O, 4·0. $C_{25}H_{34}N_5Br$, 2HBr, 1·5 H_2O requires C, 44·6; H, 5·8; N, 10·4; Br, 35·6; H_2O, 4·0%.)

N-(2-*Anilino-4-quinolylaminoethyl*)-*N′N′*-*ethyleneoxyethylenepiperazinium bromide dihydrobromide*. 2-Anilino-4-β-di(chloroethyl)aminoethylaminoquinoline dihydrochloride (2·38 g. of anhydrous salt) and morpholine (2·17 g., 5 mol.) were heated at 130° C for 2 hr. On transferring to a separating funnel with chloroform and 2N-sodium hydroxide, the solutions separated into three layers. The bottom layer of chloroform-soluble base was run off and two further extracts made. On removal of chloroform by distillation the residual base was treated with excess of aqueous 3N-hydrobromic acid and then evaporated to dryness and dissolved in hot absolute alcohol. On cooling, a gum separated which slowly crystallized. The crystalline solid was collected and dissolved in boiling absolute alcohol in which it was sparingly soluble and then separated in fluffy balls composed of minute needles, yield 0·57 g., m.p. 280° C. For analysis it was dissolved in water (0·5 ml.) and absolute alcohol (3·5 ml.) added, yield 0·48 g., m.p. 285° C of 2-*anilino-4-di(morpholinoethyl)aminoethylaminoquinoline* 3·5 *hydrobromide*. (Found: C, 46·1, 46·4; H, 5·7, 5·9; N, 11·1. $C_{29}H_{40}N_6$, 3·5 HBr requires C, 46·1; H, 5·8; N, 11·1%.)

The second pigmented oily layer insoluble in chloroform was carefully collected, after adding sodium hydroxide pellets to the aqueous layer, by decanting off the aqueous layer, and then dissolved in excess of 3N hydrobromic acid. The *quaternary salt* mentioned in the title, yield 2 g., slowly separated and was soluble initially in

methyl alcohol (20 ml.) but rapidly crystallized in needles which were then not soluble in 100 ml. of methyl alcohol, m.p. 299 to 300° C. (Found: C, 45·3; H, 5·5; N, 10·2. $C_{25}H_{32}ON_5Br$, 2HBr requires C, 45·4; H, 5·2; N, 10·6%.)

N-(2-Anilino-4-quinolylaminoethyl)-N'N'-pentamethylenepiperazinium bromide dihydrobromide. The dichloro-base dihydrochloride (2·38 g. of anhydrous salt) and piperidine (2·12 g.) were heated at 110° C for 2 hr. On adding 2N-sodium hydroxide and chloroform, three layers were obtained as in the previous experiment. The oily middle layer was collected as described above and treated with excess of 3N hydrobromic acid. The *bromide dihydrobromide* of the quaternary compound readily separated in fine needles, yield 1·8 g. It was crystallized from methyl alcohol, m.p. 300 to 302° C (efferv.). (Found: C, 46·8; H, 5·7; N, 10·6. $C_{26}H_{34}N_5Br$, 2HBr requires C, 47·4; H, 5·5; N, 10·6%.)

B. *Compounds derived from* 4-*chloroquinaldine*

4-Methylaminoquinaldine. 4-Chloroquinaldine (1·7 g., M/100) and alcoholic methylamine (13 ml., 33%, M/10) were heated in a sealed tube at 120° C for 7 hr. Excess of methylamine was distilled off and the residue made alkaline with aqueous sodium hydroxide. The *base* crystallized from ethanol in shining prisms, m.p. 238 to 239° C. (Found: C, 77·1; H, 7·4; N, 16·3. $C_{11}H_{12}N_2$ requires C, 76·7; H, 7·0; N, 16·2%.) The *hydrochloride* separated from *iso*propyl alcohol in colourless needles, m.p. 298 to 300° C. (Found: C, 54·1; H, 7·1; H_2O, 14·7. $C_{11}H_{12}N_2$, HCl, 2H_2O requires C, 54·0; H, 7·0; H_2O, 14·7%.)

4-n-Propylaminoquinaldine hydrochloride. 4-Chloroquinaldine (3·5 g., M/50) and propylamine (12·3 ml., M/10) were heated in a sealed tube at 150° C for 3 hr. The colourless recovered base (4 g.) m.p. 153 to 155° C was converted to the *hydrochloride* which crystallized from ethyl alcohol-ethyl acetate in colourless rosettes of needles, m.p. 213 to 214° C. (Found, on air-dried solid, loss at 100° C, 6·6. $C_{13}H_{16}N_2$, HCl, H_2O requires H_2O, 7·1%. On solid dried at 100° C, C, 66·4; H, 7·5; N, 11·5. $C_{13}H_{16}N_2$, HCl requires C, 65·9; H, 7·2; N, 11·8%.)

4-n-Heptylaminoquinaldine hydrochloride. 4-Chloroquinaldine (3·5 g., M/50) and *n*-heptylamine (23 g., M/5) were heated in an oil-bath at 150° C for 3 hr. The crystalline base (2·4 g.) was converted to the *hydrochloride* which crystallized as delicate rectangular leaflets from alcohol-ethyl acetate, m.p. 143 to 144° C. (Found: C, 70·4; H, 8·7; N, 9·5. $C_{17}H_{24}N_2$, HCl requires C, 69·7; H, 8·6; N, 9·6%.)

4-n-Octylaminoquinaldine hydrochloride. This salt was prepared similarly and separated from alcohol-ethyl acetate in rosettes of colourless needles, m.p. 160 to 162° C. (Found: C, 70·7; H, 9·2. $C_{18}H_{26}N_2$, 2HCl requires C, 70·5; H, 8·9%.)

4-n-Nonylaminoquinaldine hydrochloride. The components were heated at 180° C for 3 hr. and the recovered base freed from nonylamine was converted into the *hydrochloride* which crystallized from a small volume of ethyl alcohol on addition of ethyl acetate as prisms, m.p. 83° C. (Found: C, 67·1; H, 9·1; N, 8·4; H_2O, 5·2. $C_{19}H_{28}N_2$, 2HCl, H_2O requires C, 67·3; H, 9·5; N, 8·5; H_2O, 5·3%.)

4-n-Dibutylaminoquinaldine hydrochloride. 4-Chloroquinaldine (3·5 g., M/50), *n*-butylamine (12·9 g., M/10) and catalytic copper (1 g., Gattermann 1890) were heated in a sealed tube at 200° C for 5·5 hr. After basifying and removing free

dibutylamine, the residue was distilled and gave the required base (3·4 g.) as a viscous yellow oil, in 80% yield, b.p. 155 to 156° C/1 mm. The *hydrochloride* separated from ethanol-ethyl acetate in large colourless prisms, m.p. 154 to 156° C. (Found: C, 70·8; H, 9·0; N, 9·1. $C_{18}H_{26}N_2$, HCl requires C, 70·5; H, 8·9; N, 9·1%.) Under identical conditions copper bronze gave a 63% yield of base and in absence of catalyst the yield fell to 47%.

4-Morpholinoquinaldine hydrochloride. 4-Chloroquinaldine and morpholine in excess were heated at 130° C for several hours. 4-Morpholinoquinaldine base crystallized from low boiling petrol, m.p. 96° C. (Found: C, 73·1; H, 6·7. Calc. for $C_{14}H_{16}ON_2$: C, 73·6; H, 7·1%.) Meyer & Bouchet (1947) give m.p. 95° C. The *hydrochloride* was soluble in two volumes of boiling absolute alcohol and crystallized in needles, m.p. 283° C. (Found: C, 63·2; H, 6·1; N, 11·1. $C_{14}H_{16}ON_2$, HCl requires C, 63·5; H, 6·5; N, 10·6%.)

4-β-Hydroxyethylaminoquinaldine. 4-Chloroquinaldine (35·4 g., M/5) and ethanolamine (50 ml., ⅝M) were heated cautiously in an oil-bath up to 140° C. When the internal temperature began to rise rapidly, the flask was removed and placed in cold water until the vigorous reaction had subsided. The reaction was then completed by heating at 140° C for half an hour. The reaction mixture was poured into aqueous ammonia (300 ml., 2N) when *hydroxyethylaminoquinaldine* crystallized in colourless needles, m.p. 176° C, yield almost quantitative. Recrystallization from methyl ethyl ketone afforded the pure base in large prisms or tablets (40 g.) m.p. 182° C. (Found: C, 70·8; H, 6·8; N, 14·2. $C_{12}H_{14}ON_2$ requires C, 71·3; H, 6·9; N, 13·9%.) The *hydrochloride* separated from alcohol in colourless needles, m.p. 272° C. (Found: C, 60·5; H, 6·2; N, 12·3. $C_{12}H_{14}ON_2$, HCl requires C, 60·4; H, 6·3; N, 11·7%.) The *hydrobromide*, needles from alcohol had m.p. 242° C. (Found: C, 50·2; H, 5·4; N, 10·0. $C_{12}H_{14}ON_2$, HBr requires C, 50·9; H, 5·3; N, 9·9%.)

4-β-Chloroethylaminoquinaldine. 4-Hydroxyethylaminoquinaldine (20·2 g.) was added in portions to phosphorus oxychloride (60 ml.) with cooling and then the mixture was heated at 110° C for half an hour. The excess of phosphorus oxychloride was distilled off under reduced pressure and the residual caked mass was treated with several small portions of alcohol and warmed cautiously. The solution on cooling deposited the *hydrochloride* in almost quantitative yield, m.p. 272° C. (Found: C, 56·0; H, 5·5; N, 11·0. $C_{12}H_{13}N_2Cl$, HCl requires C, 56·0; H, 5·5; N, 10·9%.) The *base* was best obtained by treatment of a hot aqueous solution of the hydrochloride with concentrated aqueous ammonia. It separated from benzene in colourless cubes, m.p. 158 to 160° C. (Found: C, 65·4; H, 6·3; N, 12·9; Cl, 15·9. $C_{12}H_{13}N_2Cl$ requires C, 65·3; H, 5·9; N, 12·7; Cl, 16·1%.)

Bis-*(diethylaminoethyl)amine.* β-Chloroethyldiethylamine hydrochloride (Gough & King 1928) (57·3 g., M/3) and methanol (120 ml.) saturated at 0° C with ammonia (34 g.) were heated in an autoclave at 65 to 67° C for 2·25 hr. Water was added to the reaction mixture and the excess of ammonia and methanol removed. Strong alkali was then added and the bases thoroughly extracted with chloroform. On removing this solvent the residue was fractionally distilled to give three fractions: (i) 9 g., a colourless mobile oil, b.p. 40 to 44° C/14 to 15 mm., which when redistilled gave pure diethylaminoethylamine, b.p. 142 to 144° C (Ristenpart (1896) gives

b.p. 145° C); (ii) 14·1 g., a colourless oil, b.p. 124 to 128° C/13 to 14 mm. which on redistillation gave pure *bis*-(diethylaminoethyl)amine, b.p. 118 to 120° C/15 mm. (I.G. Farbenind. (1928) give b.p. 126 to 127° C/18 mm. whilst Soc. Chem. Ind. in Basle (1927) give b.p. 105 to 110° C/8 mm.). The identity was further confirmed by preparation of the picrate; (iii) 3·8 g., an oil, b.p. 156 to 172° C which was not examined further.

4-bis-(N-Diethylaminoethyl)aminoquinaldine trihydrobromide. 4-Chloroquinaldine (1·7 g., M/100), *bis*-(diethylaminoethyl)amine (48 g., 2·25M/100) and catalytic copper (1 g.) were heated in an oil-bath at 140° C when the temperature rose sharply to 160° C. When the vigorous reaction had subsided the external temperature was maintained at 160° C for a further 2 hr. Aqueous sodium hydroxide was then added and the dark oily layer was repeatedly washed by decantation with water to remove the excess of the aliphatic amine. The oily layer was collected in chloroform, the solvent removed and the residue distilled when a deep yellow viscous oil (2 g.), b.p. 165 to 180° C/1 mm. was obtained. N Hydrobromic acid (16 ml.) was added to this base and the resultant *trihydrobromide* separated from ethanol-ethyl acetate in colourless prisms, yield 2·1 g., m.p. 200 to 201° C. (Found: C, 41·5; H, 6·9; N, 8·8; H_2O, 7·0. $C_{22}H_{36}N_4$, 3HBr, 2·5 H_2O requires C, 41·0; H, 6·9; N, 8·7; H_2O, 7·0 %.)

4-Methylaminoethylaminoquinaldine dihydrochloride. (i) The 4-chloroethyl-base (2·2 g., M/100) and alcoholic methylamine (13 ml., 33 %, M/10) were heated in a sealed tube at 120° C for 7 hr. The oily base was isolated in the usual way and converted into the *dihydrochloride* of *4-methylaminoethylaminoquinaldine* which crystallized from ethanol-ethyl acetate in short shining needles, yield 2·2 g., m.p. 219° C, but when dried at 100° C had m.p. 250 to 251° C. (Found: C, 49·0; H, 7·4; N, 13·5; H_2O, 8·7. $C_{13}H_{17}N_3$, 2HCl, 1·5 H_2O requires C, 49·5; H, 7·0; N, 13·3; H_2O, 8·6 %.) (ii) Similar quantities of the components were heated at 150° C for 7 hr. In this case a base was isolated as a colourless solid (1·45 g.) which crystallized from ethanol in square plates, m.p. 238 to 239° C which proved to be identical with the previously described 4-methylaminoquinaldine. From the mother liquors a small quantity of the dihydrochloride of 4-methylaminoethylaminoquinaldine was obtained, m.p. 215° C.

4-Ethylaminoethylaminoquinaldine dihydrobromide. 4-Chloroethylaminoquinaldine hydrochloride (5 g.) and alcoholic ethylamine (20 ml., 33 %) were heated at 150° C. for 5 hr. The base was converted to the dihydrochloride which separated from ethanol-ethyl acetate in needles (2·1 g.) m.p. 247 to 250° C. It was purified further by conversion to the *dihydrobromide* which crystallized from spirit in needles, m.p. 248 to 250° C. (Found: C, 41·4; H, 5·7; N, 10·7; H_2O, 4·4. $C_{14}H_{19}N_3$, 2HBr, H_2O requires C, 41·1; H, 5·7; N, 10·3; H_2O, 4·7 %.)

4-n-Propylaminoethylaminoquinaldine dihydrobromide. The chloro-base (3·3 g.) and *n*-propylamine (11 ml.) were heated in a sealed tube at 150° C for 7 hr. The base was isolated in the usual way and converted into the dihydrobromide which crystallized from methanol in needles, yield 2·6 g., m.p. 266 to 270° C. (Found: C, 44·2; H, 5·9; N, 9·9. $C_{15}H_{21}N_3$, 2HBr requires C, 44·5; H, 5·7; N, 10·4 %.)

4-iso*Propylaminoethylaminoquinaldine dihydrochloride*. The chloro-base (4 g.) and *iso*propylamine (15 ml.) were heated together in a sealed tube at 120° C for 12 hr. The base was converted to the *dihydrochloride* which separated from alcohol in needles, m.p. 280° C. (Found: C, 56·8; H, 7·3; N, 13·3. $C_{15}H_{21}N_3$, 2HCl requires C, 57·0; H, 7·3; N, 13·3%.)

4-n-*Butylaminoethylaminoquinaldine dihydrobromide*. This salt was prepared in a similar way using n-butylamine. The *dihydrobromide* after several recrystallizations separated from ethanol-ethyl acetate in colourless needles, m.p. 203° C. (Found: N, 10·2. $C_{16}H_{23}N_3$, 2HBr requires N, 10·0%.)

4-tert.-*Butylaminoethylaminoquinaldine dihydrobromide*. The chloro-base (2·2 g., M/100) and *tert*.-butylamine (10·4 ml., M/10) were heated in a sealed tube at 120° C for 7 hr. The base gave, in the usual way, the *dihydrobromide* which separated from methanol-ethyl acetate in hard cubes (2·25 g.) m.p. 260 to 261° C. (Found: C, 44·1; H, 6·4; N, 10·0; H_2O, 4·3. $C_{16}H_{23}N_3$, 2HBr, H_2O requires C, 44·0; H, 6·2; N, 9·6; H_2O, 4·1%.)

4-n-*Amylaminoethylaminoquinaldine dihydrobromide*. This *dihydrobromide* was prepared as described above using n-amylamine (11·4 ml.) and was crystallized from alcohol-ethyl acetate from which it separated in fine needles, yield 1·7 g., m.p. 196 to 197·5° C. (Found: C, 44·8; H, 7·1; H_2O, 5·2. $C_{17}H_{25}N_3$, 2HBr, 1·5 H_2O requires C, 44·4; H, 6·6; H_2O, 5·9%.) From a separate experiment in which hydrochloric acid was used for salt formation, a much less soluble salt was obtained as a by-product, crystallizing from methanol in needles (1·6 g.), m.p. 312 to 313° C (decomp.). It is in all probability the trihydrochloride of the *bis*-tertiary amine (see the following two preparations).

4-n-*Hexylaminoethylaminoquinaldine dihydrobromide*. The chloro-base (2·2 g., M/100) and n-hexylamine (12 ml., M/10) were heated in an oil-bath at 120° C for 7 hr. The base was converted to the *dihydrobromide* which crystallized from alcohol on the addition of ethyl acetate in woolly needles, yield 1·9 g. Recrystallization from ethanol-ethyl acetate gave the pure salt, m.p. 200 to 201° C. (Found: C, 48·4; H, 6·7; N, 9·4. $C_{18}H_{27}N_3$, 2HBr requires C, 48·3; H, 6·5; N, 9·4%.) In a preliminary experiment where the crude base was treated with hydrochloric acid a salt very sparingly soluble in spirit was obtained which crystallized in microscopic crystals, m.p. 312° C (decomp.). Analysis showed it to be bis(4-*quinaldylaminoethyl*)*hexylamine trihydrochloride*. (Found: C, 60·8; H, 7·9; H_2O, 1·5. $C_{30}H_{39}N_5$, 3HCl, 0·5 H_2O requires C, 61·2; H, 7·4; H_2O, 1·5%.)

4-n-*Heptylaminoethylaminoquinaldine dihydrobromide*. Prepared as in the preceding using n-heptylamine (15 ml., M/10), this *dihydrobromide* separated from ethanol-ethyl acetate in woolly needles, yield 1·6 g., m.p. 199 to 200° C. (Found: C, 47·0; H, 7·2; N, 9·0; H_2O, 3·7. $C_{19}H_{29}N_3$, 2HBr, H_2O requires C, 47·6; H, 6·9; N, 8·8; H_2O, 3·7%.) In a preliminary experiment on the same scale using hydrochloric acid, a very sparingly soluble salt (0·6 g.) was obtained which separated from methyl alcohol on addition of ethyl acetate as a microcrystalline powder, m.p. 306 to 307° C identified as bis(4-*quinaldylaminoethyl*)*heptylamine trihydrochloride*. (Found: C, 61·4; H, 7·8; H_2O, 1·7. $C_{31}H_{41}N_5$, 3HCl 0·5 H_2O requires C, 61·8; H, 7·5; H_2O, 1·5%.) In another experiment, use of hydrobromic acid gave the

corresponding *trihydrobromide*, woolly needles from water, m.p. 324° C. (Found: C, 51·1; H, 6·6; N, 9·5. $C_{31}H_{41}N_5$, 3HBr requires C, 51·2; H, 6·1; N, 9·6 %.)

4-n-Octylaminoethylaminoquinaldine dihydrobromide. Prepared as in the preceding, using *n*-octylamine (16·7 ml., M/10), the excess of octylamine was distilled off under reduced pressure and the residual base was converted into the *dihydrobromide* which crystallized from ethanol-ethyl acetate in woolly needles, yield 3·3 g., m.p. 209 to 210° C. (Found: C, 50·0; H, 7·1. $C_{20}H_{31}N_3$, 2HBr requires C, 50·5; H, 7·0 %.)

4-n-Nonylaminoethylaminoquinaldine dihydrobromide. The chloro-base (2·2 g., M/100) and nonylamine (12 g., M/12) were heated in an oil-bath at 120° C for 7 hr. After removing free nonylamine the residual base was converted to the *dihydrobromide* which on repeated crystallization from ethanol finally gave fine colourless needles which showed two m.p.s 218° and 226 to 228° C. (Found: C, 49·8; H, 7·5; N, 8·0; H_2O, 3·6. $C_{21}H_{33}N_3$, 2HBr, H_2O requires C, 49·7; H, 7·4; N, 8·3; H_2O, 3·6 %.)

4-(3':5':5'-Trimethylhexylaminoethylamino)quinaldine dihydrobromide. The chloro-base (2·2 g., M/100) and 3:5:5-trimethylhexylamine (14 g.) were heated in an oil-bath at 120° C for 6 hr. The required *dihydrobromide* crystallized from ethanol in woolly needles, m.p. 246 to 248° C. (Found: C, 50·9; H, 6·9; N, 8·6. $C_{21}H_{33}N_3$, 2HBr requires C, 51·5; H, 7·2; N, 8·6 %.)

4-Anilinoethylaminoquinaldine dihydrobromide. The chloro-base (4 g.) and aniline in excess were heated in an oil-bath at 135° C for 3 hr. The required base was converted to the *dihydrobromide* which separated from ethanol in colourless needles, m.p. 252° C. (Found: C, 49·5; H, 4·7; N, 9·6. $C_{18}H_{19}N_3$, 2HBr requires C, 49·2; H, 4·8; N, 9·6 %.)

4-p-Chloroanilinoethylaminoquinaldine dihydrochloride. The chloro-base and *p*-chloroaniline in excess were heated at 160° C for 3 hr. The required *base* was sparingly soluble in boiling methyl alcohol and separated in square tablets, m.p. 197 to 199° C. (Found: C, 69·3; H, 6·2; N, 13·1. $C_{18}H_{18}N_3Cl$ requires C, 69·3; H, 5·8; N, 13·5 %.) The *dihydrochloride* crystallized from water containing a little spirit in soft needles, m.p. 170° C (efferv.). (Found: C, 54·3; H, 5·9; N, 10·6. $C_{18}H_{18}N_3Cl$, 2HCl, $\frac{3}{4}H_2O$ requires C, 54·3; H, 5·5; N, 10·6 %.) On drying at 95° C the water was not lost and at 110° C loss of hydrochloric acid and some darkening occurred as well.

4-Benzylaminoethylaminoquinaldine dihydrobromide. The parent chloro-base (2·2 g., M/100) and benzylamine (13 ml., M/10) were heated in an oil-bath at 110 to 115° C for 3 hr. The required base was freed from benzylamine by heating the recovered bases in an oil-bath at 120 to 130° C at 15 mm. pressure. The *dihydrobromide* of the required base separated from methanol-ethyl acetate in small colourless needles, m.p. 245 to 246° C. (Found: C, 47·5; H, 5·6; H_2O, 5·7. $C_{19}H_{21}N_3$, 2HBr, 1·5 H_2O requires C, 47·5; H, 5·5; H_2O, 5·6. On dried sample, found: C, 49·8; H, 5·4. $C_{19}H_{21}N_3$, 2HBr requires C, 50·3; H, 5·1 %.)

4-cycloHexylaminoethylaminoquinaldine dihydrobromide. The chloro-base (2·2 g.) and *cyclo*hexylamine (9·9 g., M/10) were heated at 115 to 120° C for 3 hr. The required base was isolated and converted to the *dihydrobromide* which separated from spirit in square plates, m.p. 246 to 248° C. (Found: C, 46·5; H, 6·1; N, 9·1;

loss at 110° C, 4·3. $C_{18}H_{25}N_3$, 2HBr, H_2O requires C, 46·6; H, 6·3; N, 9·1; H_2O, 3·9 %.)

4-β-*Phenylethylaminoethylaminoquinaldine dihydrobromide.* The chloro-base (2·2 g.) and β-phenylethylamine (7·1 g.) were heated in an oil-bath at 115 to 120° C for 3 hr. The required base was isolated and treated with hydrobromic acid and dried. The product was boiled with methanol and filtered from a sparingly soluble salt (1·0 g.), m.p. 295 to 297° C (decomp.), probably a salt containing a *bis*-tertiary base. The filtrate on concentration gave the *dihydrobromide* as hard crystals (2·5 g.) on addition of ethyl acetate. For analysis the product was finally crystallized from spirit, m.p. 248 to 250° C. (Found: C, 50·3; H, 5·6; N, 9·0. $C_{20}H_{23}N_3$, 2HBr requires C, 50·2; H, 5·4; N, 9·0 %.)

4-*Dimethylaminoethylaminoquinaldine dihydrochloride.* 4-Chloroethylaminoquin-aldine (4·0 g.) and aqueous dimethylamine (15 ml., 33 %) were heated under pressure at 120° C for 6 hr. The required base was isolated by chloroform extraction of the aqueous alkaline solution, the solvent removed, and the base converted to the *dihydrochloride*, needles from alcohol, m.p. 256° C. (Found: C, 54·4; H, 7·2; N, 13·0. $C_{14}H_{19}N_3$, 2HCl, 0·5 H_2O requires C, 54·5; H, 7·1; N, 13·5 %.)

4-*Diethylaminoethylaminoquinaldine dihydrobromide.* Prepared in the same way as the preceding, this *dihydrobromide* separated from ethanol in prisms, m.p. 272° C. (Found: C, 45·4; H, 6·1; N, 9·6. $C_{16}H_{23}N_3$, 2HBr requires C, 45·8; H, 6·0; N, 10·0 %.)

4-n-*Dipropylaminoethylaminoquinaldine dihydrochloride.* The reactants in this case were heated in an oil-bath at 110° C for 6 hr. The required base was isolated in the normal way and gave the *dihydrochloride*, needles from ethanol-ethyl acetate, m.p. 281° C. (Found: C, 59·9; H, 8·1. $C_{18}H_{27}N_3$, 2HCl requires C, 60·3; H, 8·1 %.)

4-iso*Dipropylaminoethylaminoquinaldine dihydrobromide.* In this case the re-actants diluted with absolute alcohol were heated in a sealed tube at 120° C for 12 hr. The *dihydrobromide* separated from ethanol-acetone in needles, m.p. 262° C. (Found: N, 9·1. $C_{18}H_{27}N_3$, 2HBr requires N, 9·4 %.)

4-n-*Dibutylaminoethylaminoquinaldine dihydrochloride.* The chloro-base and excess of n-dibutylamine were heated in an oil-bath at 150° C for 4 hr. The *dihydrochloride* after three crystallizations from ethanol-ethyl acetate was obtained in prisms, m.p. 254° C. (Found: C, 62·6; H, 8·8; N, 11·0. $C_{20}H_{31}N_3$, 2HCl requires C, 62·2; H, 8·6; N, 10·9 %.)

4-n-*Diamylaminoethylaminoquinaldine dihydrochloride.* Prepared in the same way as the preceding, this *dihydrochloride* crystallized from ethanol-ethyl acetate in prisms, m.p. 258° C. (Found: C, 62·5; H, 8·9; N, 10·0. $C_{22}H_{35}N_3$, 2HCl, 0·5 H_2O requires C, 62·5; H, 8·9; N, 10·0 %.)

4-n-iso*Diamylaminoethylaminoquinaldine dihydrobromide* was prepared similarly, needles, m.p. 282° C from ethanol. (Found: C, 52·1; H, 7·7; N, 8·6. $C_{22}H_{35}N_3$, 2HBr requires C, 52·5; H, 7·3; N, 8·4 %.)

4-n-*Dihexylaminoethylaminoquinaldine dihydrochloride.* The 4-chloro-base (5 g.) and n-dihexylamine (17 ml.) were heated in an oil-bath at 160 to 170° C for 1·5 hr. The required *dihydrochloride* crystallized from ethanol-ethyl acetate in prisms, m.p. 258° C. (Found: C, 64·5; H, 9·5; N, 9·5. $C_{21}H_{39}N_3$, 2HCl requires C, 65·2; H, 9·3; N, 9·5 %.)

4-n-*Diheptylaminoethylaminoquinaldine dihydrobromide*. The 4-chloro-base (2·2 g., m/100) and diheptylamine (10·65 g., m/20) were heated in an oil-bath at 160° C for 2 hr. The bases were recovered into chloroform and after removing chloroform, the excess of diheptylamine was distilled out under reduced pressure at 180° C. The residual base gave a *dihydrobromide* which crystallized from ethanol in short woolly needles, yield 2·5 g., m.p. 244 to 246° C. (Found: C, 55·9; H, 8·5. $C_{26}H_{43}N_3$, 2HBr requires C, 55·8; H, 8·1 %.)

4-bis-(*Diethylaminoethyl*)*aminoethylaminoquinaldine tetrahydrobromide*. 4-Chloro-ethylaminoquinaldine (2·2 g.) and *bis*-(diethylaminoethyl)amine (4·8 g., 2·25 m/100) were heated in an oil-bath at 120 to 130° C for 3 hr. The crystalline cake was decomposed with aqueous sodium hydroxide and the bases taken up into chloroform. The latter was then removed and the excess of the aliphatic amine removed by distillation under reduced pressure in an oil-bath at 160° C. The residue was treated with N hydrobromic acid but the salt could not be induced to crystallize. It was accordingly dissolved in water and fractionally liberated to chloroform by addition of twelve portions each of 3 ml. of N sodium hydroxide. To each of the fractions freed from chloroform, N hydrobromic acid (4 ml.) was added and the dried salt dissolved in ethanol. Fractions 4 to 9 crystallized on keeping and gave in all 2·0 g. of *tetrahydrobromide* which crystallized in small colourless prisms, m.p. 215 to 216° C, which were hygroscopic. (Found: C, 40·1; H, 6·3; N, 9·4. $C_{24}H_{41}N_5$, 4HBr requires C, 39·9; H, 6·3; N, 9·8 %.)

4-*Diethanolaminoethylaminoquinaldine*. The 4-chloro-base (4·4 g., m/50) and diethanolamine (6·3 g., 3 m/50) were heated in an oil-bath at 140 to 145° C for 3 hr. On solution in water (25 ml.) and addition of aqueous sodium hydroxide a base was precipitated in solid form. It was converted to the *dihydrobromide* by addition of N hydrobromic acid (40 ml.). This salt was crystallized from aqueous alcohol-acetone from which it separated in needles, yield 7·2 g. When air-dried it partially melts below 130° C but when dried at 100° C it may show partial melting between 170° and 180° C and then finally melt at 207° C. (Found on dried salt: C, 42·9; H, 5·6; N, 9·6. $C_{16}H_{21}O_2N_3$, 2HBr requires C, 42·6; H, 5·5; N, 9·3 %.) The *base* prepared from the pure dihydrobromide crystallized from ethanol-ethyl acetate in prisms, m.p. 118° C. (Found: C, 71·0; H, 7·6; N, 15·7. $C_{16}H_{21}O_2N_3$ requires C, 70·8; H, 7·7; N, 15·5 %.)

4-*Di*(*chloroethyl*)*aminoethylaminoquinaldine dihydrochloride*. 4-Diethanolamino-ethylaminoquinaldine (2·8 g.) and phosphorus oxychloride (20 ml.) were heated in an oil-bath at 100 to 120° C for 2 hr. The excess of oxychloride was removed and the viscous mass which was left was moistened with alcohol and slightly warmed; it was thus converted into the crystalline *dihydrochloride* which was collected and recrystallized from methanol containing a drop of water, from which it separated in colourless needles, yield 2·5 g., m.p. 240° C. (Found: C, 48·3; H, 6·3; N, 10·4. $C_{16}H_{21}N_3Cl_2$, 2HCl requires C, 48·1; H, 5·8; N, 10·5 %.)

4-*Piperidinoethylaminoquinaldine dihydrochloride*. The parent 4-chloroethyl-aminoquinaldine base (3·55 g.) and excess of piperidine were gently refluxed for 4 hr. After recovering the mixture of bases and removing the piperidine the residual base was extracted with boiling ligroin which on concentration gave the

required base, m.p. 92° C. The *dihydrochloride* was obtained pure after two crystallizations from *iso*propyl alcohol from which it separated in prisms, m.p. 268° C. (Found: C, 58·7; H, 7·3; N, 11·7. $C_{17}H_{23}N_3$, 2HCl, 0·5 H_2O requires C, 58·2; H, 7·4; N, 11·9%.)

4-Morpholinoethylaminoquinaldine. The parent 4-chloro-base (2 g.) and morpholine (10 ml.) were heated in an oil-bath at 135° C for 3 hr. The crude base was obtained in the usual way and after preliminary crystallization as a hydrochloride which could not be obtained quite pure was again recovered as *base* which separated from benzene-ligroin in prisms, m.p. 120° C. (Found: C, 71·0; H, 7·6; N, 15·7. $C_{16}H_{21}ON_3$ requires C, 70·8; H, 7·7; N, 15·5%.) The *dihydrochloride* prepared from the pure base separated from *iso*propyl alcohol in colourless needles, m.p. 268° C. (Found: C, 53·2; H, 7·0; N, 11·6; H_2O, 6·0. $C_{16}H_{21}ON_3$, 2HCl, H_2O requires C, 53·0; H, 6·9; N, 11·6; H_2O, 5·0%.)

4-Dibenzylaminoethylaminoquinaldine dihydrobromide. The chloro-base (4 g.) and excess of dibenzylamine were heated at 135° C. for 3 hr. The *dihydrobromide* prepared in the usual way separated from ethanol in needles, m.p. 196° C. (Found: C, 56·7; H, 5·4; N, 7·9. $C_{26}H_{31}N_3$, 2HBr requires C, 57·0; H, 6·0; N, 7·7%.)

Bis-(4-quinaldylaminoethyl)methylamine. 4-Chloroethylaminoquinaldine (4·4 g., M/50), alcoholic methylamine (2 ml., 33%, 3M/200) and alcohol (8 ml.) were heated in a sealed tube at 130° C for 7 hr. The required base was isolated as an oil which was neutralized to Congo-paper with N-hydrochloric acid (33 ml.). The solution was evaporated to dryness and the *trihydrochloride* crystallized from aqueous alcohol from which it separated in short colourless needles (2·3 g.) m.p. 302 to 304° C. (Found: C, 51·8; H, 7·1; N, 11·9; H_2O, 12·4. $C_{25}H_{29}N_5$, 3HCl, 4H_2O requires C, 51·7; H, 6·9; N, 12·1; H_2O, 12·4%.) The *base* prepared from the pure salt crystallized from benzene in hard cubes melting first at 142° C and after resolidifying at 171 to 172° C. For analysis the base was dried to constant weight at 100° C. (Found: C, 74·6; H, 7·2; N, 17·2. $C_{25}H_{29}N_5$ requires C, 75·1; H, 7·3; N, 17·5%.)

Bis-(N-4-quinaldylaminoethyl)γγ'-dipiperidyl. The 4-chloro-base (5·5 g.) and sublimed γγ'-dipiperidyl (2·75 g.) were heated together in phenol at 140° C for 3 hr. The *base* resulting from this reaction was found to be sparingly soluble in chloroform or alcohol but was crystallized from pyridine from which it separated in small prisms, m.p. 235° C. (Found: C, 76·0; H, 8·3; N, 15·5. $C_{34}H_{44}N_6$ requires C, 76·1; H, 8·2; N, 15·7%.) The *tetrahydrochloride* m.p. 326° C. (decomp) crystallized from water in prisms. (Found: C, 51·3; H, 7·1; N, 10·6; H_2O, 13·1. $C_{34}H_{42}N_6$, 4HCl, 6H_2O requires C, 51·6; H, 7·6; N, 10·6; H_2O, 13·7%.)

N-(4-Quinaldylaminoethyl)-N'-n-hexylpiperazine trihydrochloride. 4-Di(chloroethyl)aminoethylaminoquinaldine dihydrochloride (1·6 g.) and *n*-hexylamine (3·6 ml.) were heated on a boiling water-bath for 4 hr. The required base was freed from hexylamine and converted to the *trihydrochloride* which separated as a spongy mass of fine needles (1·3 g.), m.p. 284° C (decomp.). (Found: C, 57·0; H, 8·5; N, 11·7. $C_{22}H_{34}N_4$, 3HCl requires C, 57·0; H, 8·0; N, 12·1%.)

N-(4-Quinaldylaminoethyl)-N'N'-diethylpiperazinium bromide dihydrobromide. 4-Di(chloroethyl)aminoethylaminoquinaldine dihydrochloride (2·5 g., 3M/400) and diethylamine (8·76 g., 3M/25) were heated in a sealed tube at 120° C for 7 hr. The

contents of the tube were taken to dryness, treated with 2N-sodium hydroxide and extracted with chloroform. The chloroform extract was examined but only contained a trace of basic oil. The parent alkaline solution was treated with solid sodium hydroxide when an oily layer separated. This was carefully collected, treated with N hydrobromic acid (30 ml.) and the solution evaporated to dryness. The *bromide dihydrobromide* so obtained separated from methanol in colourless needles (2·1 g.), m.p. 245 to 246° C. On recrystallization the pure salt was obtained showing two melting points, 246° and 254 to 255° C. (Found: C, 41·4; H, 5·9; N, 9·6; H_2O, 1·7. $C_{20}H_{31}N_4Br$, 2HBr, 0·5 H_2O requires C, 41·5; H, 5·9; N, 9·9; H_2O, 1·6%.)

N-(4-Quinaldylaminoethyl)-N'N'-pentamethylenepiperazinium bromide dihydrobromide. The dihydrochloride of the dichloro-base (2·4 g.) and piperidine (10 ml.) were heated on the boiling water-bath for 5 hr. The excess of piperidine was immediately removed by heating under diminished pressure, water and 2N sodium hydroxide added to the residue and a small quantity of basic oil removed by chloroform. Solid sodium hydroxide was then added to the alkaline solution and the brown oil which separated was carefully collected and treated with an excess of hydrobromic acid. The *bromide dihydrobromide* crystallized from methanol containing a trace of water in needles (2·86 g.). For analysis it was crystallized from 8 volumes of spirit containing a few drops of 3N hydrobromic acid and separated in clusters of fine needles, m.p. 293 to 294° C. (Found: C, 39·1; H, 6·5; N, 8·5; H_2O, 10·4. $C_{21}H_{31}N_4Br$, 2HBr, 4H_2O requires C, 38·6; H, 6·3; N, 8·6; H_2O, 11·0%.)

N-(4-Quinaldylaminoethyl)-N'N'-ethyleneoxyethylenepiperazinium bromide dihydrobromide. The dichloro-base dihydrochloride (4 g.) was regenerated to base which latter was then heated with morpholine on the water-bath for 5 hr. Almost immediately a colourless crystalline material began to separate which was eventually collected, washed with ether, dissolved in water and treated with concentrated aqueous sodium hydroxide. The solid which was precipitated was carefully collected and treated with excess of hydrobromic acid. The *bromide dihydrobromide* separated from aqueous alcohol in well-formed colourless prisms, m.p. 283° C. (Found on macro-analysis, loss at 110° C, 10·4. $C_{20}H_{29}N_4OBr$, 2HBr, 4H_2O requires H_2O, 11·0%. On micro-analysis, loss at 100°, 6·3. $C_{20}H_{29}N_4OBr$, 2HBr, 4H_2O losing 2H_2O requires H_2O, 5·8%. On this dried sample, found: C, 39·0; H, 6·1; N, 9·7. $C_{20}H_{29}N_4OBr$, 2HBr, 2H_2O requires C, 38·8; H, 5·7; N, 9·1%.)

REFERENCES

Anderson, H. H. & Chuan, T. T. K. 1944 *Amer. J. Trop. Med.* 24, 367.
Baeyer, A. & Bloem, F. 1882 *Ber. dtsch. chem. Ges.* 15, 2152.
Bleier, L. 1899 *Ber. dtsch. chem. Ges.* 32, 1829.
Boeck, W. C. & Drbohlav, J. 1925 *Amer. J. Hyg.* 51, 371.
Brackett, S. & Bliznick, A. 1947 *J. Parasit.* 33, 154.

Brooker, L. G. S. & Smith, L. A. 1937 *J. Amer. Chem. Soc.* **59**, 72.

Buchmann, F. J. & Hamilton, C. S. 1942 *J. Amer. Chem. Soc.* **64**, 1357.

Child, R. & Pyman, F. L. 1929 *J. Chem. Soc.* p. 2010.

Curd, F. H. S., Raison, C. G. & Rose, F. L. 1947 *J. Chem. Soc.* p. 905.

Dobell, C. 1947 *Ann. Soc. belge Med. trop. Liber Jubilaris J. Rodhain*, p. 201.

Dobell, C. & Laidlaw, P. P. 1926 *Parasitology*, **18**, 206.

Dziewonski, K. & Dymek, W. 1937 *Chem. Zbl.* **1**, 1153.

Gattermann, L. 1890 *Ber. dtsch. chem. Ges.* **23**, 1219.

Goodson, J. A., Goodwin, L. G., Gorvin, J. H., Goss, M. D., Kirby, K. S., Lock, J. A., Neal, R. A., Sharp, T. M. & Solomon, W. 1948*a* *Brit. J. Pharmacol.* **3**, 49.

Goodson, J. A., Goodwin, L. G., Gorvin, J. H., Goss, M. D., Kirby, K. S., Lock, J. A., Neal, R. A., Sharp, T. M. & Solomon, W. 1948*b* *Brit. J. Pharmacol.* **3**, 62.

Goodwin, L. G., Hoare, C. A. & Sharp, T. M. 1948 *Brit. J. Pharmacol.* **3**, 44.

Gough, G. A. C. & King, H. 1928 *J. Chem. Soc.* p. 2436.

Hargreaves, W. H. 1945 *Lancet*, no. 249, p. 68.

I.G. Farbenind. 1928 British Patent, 292,615.

Jones, W. R. 1946 *Ann. Trop. Med. Parasit.* **40**, 130.

Jones, W. R. 1947 *Brit. J. Pharmacol.* **2**, 217.

Laidlaw, P. P., Dobell, C. & Bishop, A. 1928 *Parasitology*, **20**, 207.

Meyer, A. & Bouchet, G. 1947 *C.R. Acad. Sci., Paris*, **225**, 63.

Pyman, F. L. 1937 *Chem. Ind.* **56**, 789.

Pyman, F. L. & Wenyon, C. M. 1917 *J. Pharmacol.* **10**, 237.

Rawson, G. W. & Hitchcock, D. J. 1947 *J. Parasit.* **33**, 19.

Ristenpart, E. 1896 *Ber. dtsch. chem. Ges.* **29**, 2526.

Rügheimer, L. 1884 *Ber. dtsch. chem. Ges.* **17**, 737.

Rügheimer, L. 1885 *Ber. dtsch. chem. Ges.* **18**, 2978.

Schönhöfer, F. 1948 *Fiat Review of German Science*, 1939–1946, *Chemotherapy*, p. 85.

Soc. Chem. Ind. in Basle 1927 British Patent, 310,074.

Vedder, E. B. 1911 *Bull. Manila Med. Soc.* **3**, 48.

Wright, H. D. 1933 *J. Path. Bact.* **37**, 257.

Experimental studies on amphibian oocyte nuclei

I. Investigation of the structure of the nuclear membrane by means of the electron microscope

By H. G. Callan, *Animal Breeding and Genetics Research Organization,*
Agricultural Research Council, Institute of Animal Genetics, Edinburgh

and S. G. Tomlin, *Wheatstone Physics Laboratory,*
King's College, London

[Plates 18 to 19]

The membrane surrounding the nucleus of the oocytes of two species of amphibian is shown to consist of two structures, an outer porous layer and an inner layer which is apparently continuous. The porous layer is about twice as thick as the inner layer, the dried membrane as a whole having a thickness of approximately 500 Å. The pores are of regular size and arrangement in the outer layer; pore diameter is approximately 400 Å, the separation distance between pore centres 1000 Å. Both layers consist of relatively insoluble protein materials, the porous layer also containing some lipoid.

1. Introduction

Although the nuclear membrane is a characteristic feature in the structural organization of the vast majority of plant and animal cells, comparatively little attention has been paid to its nature and properties. The reasons for this neglect are not hard to seek; most cells are small objects, nuclei still smaller. Physiologists have devoted a great deal of attention to the plasma membrane and the cell cytoplasm, considerably less to the cell nucleus. Cytologists have devoted more of their attention to the cell nucleus, less to the cell cytoplasm; the nuclear membrane, which does not normally undergo spectacular changes during the life of the cell, is generally dismissed as either present or absent according to the stage in the cycle of cell division which is under examination.

As recently as 1937, Pischinger has denied the existence of an organized membrane surrounding the nucleus. He regards the apparent membrane as a physical interface, and nothing more. Yet the simplest experiments in micro-dissection of a wide variety of cells supply convincing evidence that the nuclear membrane is an organized structure with detectable mechanical strength. If, by means of a micro-manipulator, one attempts to puncture a cell such as an *Amoeba* or salivary gland cell of *Drosophila*, the needle meets resistance at the plasma membrane; having overcome this resistance, the needle moves relatively freely in the cell interior until it meets the nuclear membrane. Here further resistance is encountered; once the nuclear membrane has been punctured, the needle again moves relatively freely inside the nucleus.

Baud (1948) has reviewed the evidence for the existence of a nuclear membrane in all cells containing nuclei. He points out that whereas in certain cells the membrane is clearly visible in the living state, in other cells its existence can only be inferred by indirect means.

Unlike most biological membranes, the nuclear membrane is a wholly intra-cellular structure which divides the cell into two environments differing from one another both chemically and physically. Across this membrane there are movements of metabolites, of special interest being those associated with gene activity. It is therefore evident that the structure and properties of the nuclear membrane merit experimental study.

The small size of most nuclei has placed them outside the direct reach of the experimenter by techniques which do not involve micro-manipulation. There are, however, various animal cells, themselves of large size, which contain nuclei sufficiently large to be dealt with freehand. The nuclei (germinal vesicles) of maturing amphibian oocytes may attain diameters approaching 1 mm. and are particularly suitable for study. These nuclei have been used as material for a wide variety of investigations by Brachet (1938, 1940), Duryee (1937, 1938, 1941, 1942), Gersch (1940) and Waddington (1938). Although these authors have made occasional references to the membrane, none have been concerned with a study of its structure.

The present paper is one of a series which will describe experiments with oocyte nuclei of amphibians. The aim of the experiments is to exploit these giant nuclei systematically in order to obtain information which, with certain reservations, may have general application to nuclei less suited to experimental study.

2. MATERIAL AND GENERAL METHODS

Two species of Amphibia have been used for the present study, *Triturus cristatus cristatus* and *Xenopus laevis*. In most experiments the newt or toad was anaesthetized with ether, the abdominal body wall cut longitudinally for about 1 cm. to one side of the mid-ventral line, and part of the ovary excised. The wound was subsequently sutured with two or three stitches of gut and the animal allowed to recover. Regeneration of a continuous body wall normally occurs within 1 month from the time of operation.

In early experiments the ovarian fragments were placed in amphibian Ringer solutions; the oocytes distend in such media and rapidly undergo irreversible damage; they maintain good condition for somewhat longer periods in Ringer solutions containing 0·5 % gelatine, but ultimately it was found preferable to dispense with aqueous solutions entirely. Excised fragments were placed in dry embryo-cups covered by glass plates sealed with liquid paraffin, or they were totally immersed in liquid paraffin. Where slight contamination with paraffin is of no disadvantage, the latter method is excellent; the majority of the oocytes remain bathed in the body fluid of the animal from which they were removed and they do not come into direct contact with the paraffin, which, nevertheless, seals the pre-paration from direct contact with the atmosphere. Although fresh material was used throughout most of these experiments, it may be worth while to record that

ovarian fragments immersed in paraffin and held at a temperature of 3 to 6° C maintain splendid condition for over 24 hr.

The young oocytes are transparent and their nuclei are readily visible as 'granular' spheres, the granulations being in fact large numbers of nucleoli which lie adjacent to the inner surface of the nuclear membrane. In older oocytes the yolky cytoplasm is opaque and the nuclei are not visible from the outside. In order to isolate nuclei from full-grown oocytes, the following procedure is adopted. A fragment of ovary bearing a few large oocytes is rinsed several times in the medium intended for isolation and then placed, in its container, on the stage of a low-power binocular microscope (magnification × 16). The stalk of attachment of a full-grown oocyte is now seized with watchmakers' forceps and the follicular envelope surrounding the oocyte punctured with a shallow, almost tangential, prick from a sharp mounted needle. In the case of newt oocytes, provided the isolating medium does not coagulate the yolky cytoplasm, the latter flows out in a steady stream from the puncture when slight pressure is applied by the needle to the oocyte's surface. It is generally possible to see the position of the nucleus, which stands out as a swelling on the extruded ribbon of cytoplasm. The nucleus can be removed from its adherent cytoplasm by the gentle pumping action of a pipette whose orifice is somewhat wider than the cross-section of the nucleus. The orifice of the pipette should have been smoothed off in a small gas flame prior to use.

In the case of *Xenopus* oocytes, the position of the nucleus can be judged from the outside; it lies immediately below the surface at the pigmented pole. The yolky cytoplasm of *Xenopus* oocytes is very much stiffer in consistency than that of the newt. With practice it is possible to tear the follicular envelope immediately overlying the nucleus, in which case the nucleus often pops out from the oocyte already free from adhering cytoplasm. More frequently, because of an awkward disposition of the stalk with respect to the pigmented pole, the oocyte must be punctured elsewhere. In this case the oocyte must be torn open, exerting as little pressure on the stiff cytoplasm as is possible. If a *Xenopus* oocyte is punctured and squeezed, the nuclear sap generally bursts out through a small hole in the nuclear membrane and the nuclear isolation fails.

The nuclei from full-grown oocytes of *Triturus* and *Xenopus* contain large numbers of nucleoli which form a small central mass surrounding the chromosomes. In *Xenopus*, and in *Triturus* nuclei which are not fully mature, there are also peripheral nucleoli adjacent the nuclear membrane. Lying between the group of central nucleoli and the membrane is a wide layer of nuclear sap. *Triturus* nuclei are spheres or slightly prolate spheroids with a regular outline. *Xenopus* nuclei, on the other hand, though being spherical, have a very uneven outline. The membrane is sacculated, and if the nucleus distends, as it does if isolated into a non-coagulating colloid-free medium, the sacculations at first project stiffly from a taut sphere; if the nucleus distends to an extreme degree the areas of membrane between sacculations expand, taking up the strain, until this nucleus too becomes a sphere of even outline.

For most experimental purposes the easily isolated, larger and smooth outlined nuclei of *Triturus* are preferable to those of *Xenopus*. However, they can only be

obtained for limited periods during the year, unlike those of *Xenopus*, which are available throughout the year. Most of the work with the electron microscope has been carried out on *Xenopus* nuclei, the *Triturus* material being in the main reserved for experiments where measurements of surface area were involved. However, sufficient *Triturus* material was studied to show that it does not differ materially from *Xenopus* so far as the fine structure of the nuclear membrane is concerned.

Electron micrographs were obtained using a Siemens electron microscope operating in all cases at 52kV and at a magnification of 13,000 times. The plates used were either Barnet Process or Ilford Rapid Process, Experimental, and they were developed in undiluted Kodak D.8 developer. Exposure times were 2 to 4 sec.

Most of the preparations were mounted on conventional copper grids with or without supporting films of Formvar (polyvinyl formal). A few were mounted on platinum disks and a further few on stainless steel grids in order to avoid the effects of metal ions during chemical treatment of the mounted specimens.

3. Observations

The first isolations of nuclei (*Triturus*) for examination of membranes in the electron microscope were made in unbuffered solutions of sodium chloride at 0·8 % concentration. This solution is specifically *not* an adequate physiological medium for nuclear isolation. The nuclei distend to about $1\frac{1}{2}$ times their original diameter owing to the absence from the medium of a non-penetrating colloid balancing the osmotic pressure of the nuclear sap colloids. This is no disadvantage for the examination of the membrane, since one thereby starts with a nucleus larger than life and whose membrane is stretched free from wrinkles and minor irregularities of contour. So far as electrolytes alone are concerned, sodium chloride solution is not a suitable medium for the preservation of nuclear sap and nucleoli; however, the fine structure of the nuclear membrane seems to be as well preserved in sodium chloride solution as in more complex media designed to maintain the nuclear colloids in life-like condition.

Nuclei were allowed to distend after isolation, and their membranes were then ruptured and stretched over the central squares of copper mounting grids by means of fine-pointed tungsten needles. These needles were made by the method recommended by Pantin (1946). Fortunately, the membrane is adhesive when punctured; it is therefore not unduly difficult to obtain a flat preparation lying with its inner surface directed upwards and its outer surface in contact with the grid. Some of the preparations made in this way were washed thoroughly in sodium chloride solution followed by distilled water and dried off in a vacuum desiccator without prior fixation. Such preparations were found to be relatively free from foreign matter when examined in the electron microscope; they were, however, very uneven in texture, a condition apparently resulting from stresses set up during desiccation. They gave the appearance of an irregular meshwork supporting hosts of small annuli; in favourable patches which by accident had been exposed to less severe stresses during desiccation, as happens, for example, in areas circumscribed by folded membrane, the annuli were replaced by a continuous sheet bearing pores of approximately the same size as those within annuli.

Later preparations, mostly of *Xenopus*, mounted direct on grids, were handled in a similar manner, but slight modifications in procedure were adopted. *Xenopus* nuclei were isolated in distilled water rather than in sodium chloride in order that the membrane should be stretched free from most sacculations before being mounted. When the nuclei are isolated in distilled water they stretch beyond the normal limit of 1½ times the initial diameter (this happens also to be the limit of elastic expansion below which no permanent deformation is produced). If left unchecked, the nuclei will continue to distend until they burst, the nucleoli and other nuclear contents all going into a homogeneous solution a few seconds after isolation and developing an excessively high osmotic pressure in the process. However, the nuclei were punctured and the membranes stretched over grids before extreme distension had occurred.

Membranes isolated in distilled water were for the most part fixed in 2 or 0·1 % phosphotungstic acid or in osmic acid vapour. A few membranes were washed in 10 % sodium chloride solution after mounting and before fixation, the aim being to effect the complete removal of contaminating cytoplasmic and nuclear sap materials, following the principles advocated by Bensley (1938). Clean preparations were obtained in this way without disturbance of the membrane structure; the insolubility of the nuclear membrane has already been noted by Bensley.

Preparations made as described above enabled one component of the membrane structure to be identified with certainty. This component is a sheet of material with pores evenly spaced out over its area. It will be referred to as the *a*-layer. In these preparations the pore diameter is approximately 300 Å, the distance between pore centres 800 Å. A typical area is shown in figure 1, plate 18.

It is of the utmost importance to recognize that the methods of preparation of biological structures for examination in the electron microscope are likely to give rise to artifacts of various kinds. There are four stages at which such artifacts can occur: (*a*) in the course of manipulating the fresh material; (*b*) in the process of fixation; (*c*) as a result of desiccation, which is a prerequisite for examination in the electron microscope; and (*d*) as a result of the heating of the preparation while it is lying in the electron beam.

Artefacts originating from sources (*a*) and (*b*) can be guarded against by comparing the results obtained with a variety of experimental procedures. The structure of the *a*-layer is repeatedly observed whether the membranes are isolated in distilled water or in physiological salines, and, within reasonable limits, it does not vary with the time interval between isolation and fixation. Moreover, the same structure is visible whether the membranes are fixed in phosphotungstic acid, osmic acid, trichloracetic acid or alcohol (provided in the latter case, as will be seen later, extraction of lipoids is prevented). As has already been stated, the structure is sufficiently stable to be readily recognizable even in unfixed preparations. A further argument against the structure of the *a*-layer being an artefact lies in the observation that pore size and distribution are of great regularity. Were this structure an artefact one would expect considerable variation both in pore size and distribution, especially when preparations made by different techniques are compared. This is decidedly not the case.

Artefacts arising from source (c) are to some extent linked with those from source (b); the less water a structure contains, the more likely it is to fix and dry without distortion. Compared with other biological structures, the nuclear membrane is a strikingly dense object, and it shows no evidence of hydrating on isolation. However, it is probable that the dimensions of pore diameter and pore separation distance already quoted are somewhat underestimates of the dimensions in life, since nuclear membranes which are not supported undergo shrinkage on fixation and drying. They should therefore be accepted as minimal figures only.

Artefacts arising from source (d) can be checked by comparing the structure of a preparation when first examined in the microscope with that after a lengthy exposure. During examination in the electron beam the sheets of membrane often contract and split and curl upon themselves. This gross mechanical effect does not disturb the fine structure of the membrane except in so far as it may add to an overall underestimate of dimensions. By accident of fixation, which we are not able to define precisely, some preparations do show a readily recognizable artefact. Instead of a porous membrane these preparations consist of sheets of annuli which are strung together in irregular lines (figure 3, plate 19). It is immediately evident on examination of photographs of preparations such as these that neighbouring groups of annuli fit together as in a jigsaw puzzle; they are, in fact, derived by the mechanical breakdown of the porous sheets which are seen in the majority of preparations. A noteworthy point established by 'annular' preparations is that the regions of mechanical weakness in the porous membrane lie between pores rather than across them. It may be that the rims of the pores are built up above the general level of the membrane; indeed, some photographs give a distinct impression that this is the case.

A question is raised by the 'annular' preparations; the groups of annuli are often well separated from their neighbours and yet remain in the form of a sheet; they must therefore be supported by a further structure. This second component is a continuous membrane, which will in future be referred to as the b-layer. It was first seen in some of the early preparations of *Triturus* which had been made from oocytes excised 24 hr. before isolation of the nuclei and held at room temperature. Preparations of such nuclei showed no traces of the a-layer; instead, they showed a continuous membrane covered with unidentifiable debris. The a-layer had evidently broken down by autolysis during the interval between excision of the oocytes and the mounting of the membrane. Similarly, other early preparations which had had a very extended washing in distilled water after mounting and before fixation showed no signs of the structure of the a-layer but again suggested the existence of a continuous layer supporting debris.

It was later found that the a-layer could be destroyed at will. Nuclei were isolated in 0·2 M-potassium phosphate solution containing magnesium phosphate at saturation. At pH ranging from 6·6 to 6·8 (determined by glass electrode measurements) the nuclear sap components have a very life-like appearance apart from the distension which normally occurs in colloid-free media. The external surfaces of nuclear membranes are remarkably sticky when isolations are made in phosphate. It was felt that this property could be used to advantage in a mounting technique

not requiring the use of needles. Nuclei were isolated in phosphate solutions and pipetted on to cover-slips previously coated with films of Formvar. On removal of the bulk of the liquid, the nuclei flatten down and their lower surfaces become rigidly attached to the Formvar. If a single drop of liquid is now released from a pipette held vertically above the flattened nucleus, in most instances the membrane ruptures in a ring around the area adhering to the Formvar leaving thereby a flat area of membrane whose inner surface is directed upwards. The membrane can be fixed and the film plus membrane subsequently separated from the glass cover-slip and together mounted on a copper grid ready for examination. Technical details of this mounting technique, which has been developed for the study of tissue culture cells, are given by Martin & Tomlin (1950). Preparations made in this way show the *b*-layer lying on the Formvar film; of the *a*-layer there is no trace, but irregular scattered debris probably represents the products of its breakdown. The disruption of the *a*-layer is effected specifically by the phosphate ion and it occurs over a wide pH range. In order to study the structure of the *b*-layer, membranes from nuclei isolated in phosphate were stretched directly over copper grids without a supporting Formvar film. Figure 4, plate 19, is a photograph of a preparation of this kind. It shows the *b*-layer, without evident fine structure, bearing debris derived from the *a*-layer.

The *a*-layer can also be disrupted mechanically. It is a property of the nuclear membrane to lie at an air-water interface in preference to total immersion in water. Gross (1916) made a similar observation when working with the nuclei of molluscan oocytes. Though this property is particularly striking when nuclei are isolated in phosphate solutions, the direct effects of which have just been noted, it also occurs to a lesser extent with nuclei isolated in distilled water or in chloride solutions. It is possible to pick up the part of a membrane lying in the air-water interface directly on a grid or on a Formvar which can subsequently be mounted on a grid. Preparations made in this way show the *b*-layer covered with debris from the *a*-layer, but very little trace of the latter's structure remains. With this effect in mind, it is small wonder that the first unfixed membranes should have been uneven in texture, since on drying out they too were exposed to surface-tension forces. It is remarkable that any trace of the structure of the *a*-layer should have remained at all!

It was now necessary to determine the topographical relationship of the *a*- and *b*-layers. Proofs negative and positive were obtained that the *a*-layer lies on the outside of the nucleus. Nuclei of *Xenopus* were isolated in distilled water and pipetted on to cover-slips bearing a Formvar film. The same technique was adopted as with nuclei isolated in phosphate solution; water was abstracted and the nuclei flattened, and then a drop of water pipetted from above directly over the area of membrane lying in the air-water interface. The membranes in distilled water are not noticeably adhesive, and it is only in a small proportion of cases that a piece of membrane is left adhering to the Formvar after rupture of the nucleus has occurred. In the few successful cases the membrane adhering to Formvar (which was not exposed to the surface tension forces at the interface prior to fixation) was subsequently fixed in 0·1 % phosphotungstic acid solution and, after washing in distilled

water, mounted together with the Formvar film on a copper grid. A photograph of a typical preparation of this kind is shown in figure 2, plate 18. It will be seen that the a-layer is beautifully preserved in all its regularity; photographic contrast is, however, less than in preparations mounted direct on grids owing to the presence of the continuous film of Formvar as well as the nuclear membrane itself. Pore diameter is 500 Å, the distance between pore centres 1300 Å. It is likely that these dimensions are somewhat larger than those existing in life, since the membranes were expanded by osmotic pressure before fixation to the Formvar.

Preparations made in the way just described were shadowed *in vacuo* with palladium (Williams & Wyckoff 1944). If the a-layer lies free at the surface exposed to shadowing, then the pore rims should cast shadows over the pores themselves. However, this was not the condition which we observed in the photographs. The effect of shadowing was merely to decrease still further the photographic contrast, though odd fragments of debris lying on the upper surface threw shadows in the normal way. This negative result indicates that the a-layer is not lying exposed at the surface; it is covered by the uniform b-layer (see figure 6a). This means that in the case of the intact nucleus the a-layer lies outside, the b-layer inside.
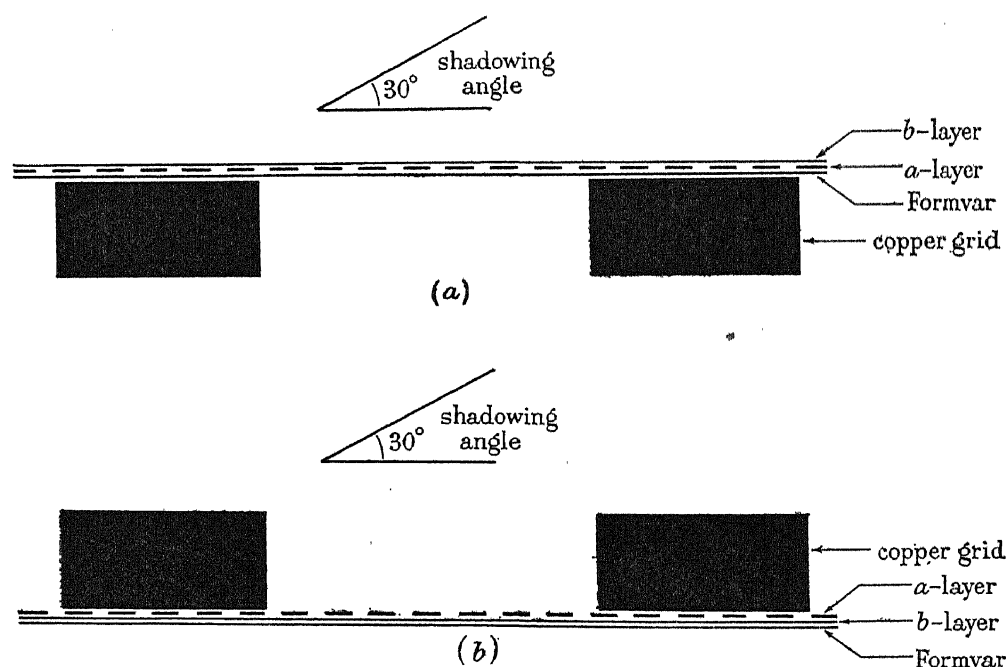
FIGURE 6. The disposition of the two layers of the nuclear membrane in relation to the angle of incidence of palladium particles in two kinds of shadowing arrangement. (a) The membrane lying with the a-layer in contact with Formvar, the Formvar being supported on a copper grid. Being overlaid with the b-layer, the rims of the pores of the a-layer cast no shadows. (b) The membrane lying with the a-layer attached to a copper grid, the b-layer being backed by Formvar. Where the a-layer is not itself in the shadows of the grid, the rims of its pores cast shadows.

Proof positive of the topographical relationship was also established. *Xenopus* nuclei were isolated in distilled water and mounted direct on copper grids by means

of tungsten needles. They were then fixed in 0·1 % phosphotungstic acid solution. In these preparations the outer layer of the membrane is in contact with the grid, the inner layer directed upwards. Free sheets of Formvar were now prepared on the air-water interface and the grids carrying membranes brought up to the surface, so that the inner membrane of the nucleus was overlaid by Formvar. After drying, the preparations were inverted, set up *in vacuo* and shadowed with palladium. In these preparations, as will be evident from figure 6*b*, the outer layer of the nuclear membrane is directly exposed to the metallic particles where it is not itself in the shadows cast by the ribs of the mounting grid. Photographs such as that shown on figure 5, plate 19, clearly demonstrate that it is the *a*-layer which is lying free at the surface; the layer is badly broken up, but sufficient remains to establish the point at issue. Where the *a*-layer is interrupted, the adjoining parts cast shadows on to the *b*-layer. From the width of the shadow and knowing the shadowing angle (30°) we can make an approximate estimate of the thickness of the *a*-layer. It is roughly 300 Å thick. The *b*-layer is itself lying free and torn in places. Judging from the width of the shadows which it casts on to the supporting Formvar film below, its thickness is roughly 150 Å.

A further proof of the topographical relationship was obtained by means of replicas of the surfaces of whole fixed nuclei. The nuclei were fixed with trichloracetic acid and dehydrated with alcohol. One of these was then placed on a clean cover-slip. A second cover-slip coated with a film of collodion in amyl acetate solution was then brought down carefully on to the nucleus before the collodion film had dried out and before the nucleus had dried fast to the lower cover-slip. On lifting the coated cover-slip the nucleus adhered to the collodion which was then allowed to dry out. After drying, a drop of 0·1 % trypsin solution was placed over the nucleus, the cover-slip inverted and sealed over a depression slide with paraffin wax. The preparation was incubated at 37° C for 12 hr. It was subsequently opened, the debris washed away and the portion of collodion film bearing the replica of the nucleus was mounted on a copper grid and shadowed with palladium. Photographs of such preparations gave further indubitable evidence that the *a*-layer is on the outside of the nucleus.

It was thought desirable to confirm the estimates of thickness made from shadow-length measurements by an independent method. The optical method of multiple beam interferometry described and applied to many physical problems by Tolansky (1948) was used to measure the total thickness of the membrane, the procedure being as follows. *Xenopus* nuclei were isolated in saline and transferred to carefully cleaned cover-slips. The excess saline was now pipetted off and the nuclei ruptured, leaving the lower halves stuck to the cover-slips as earlier described. The membranes were now fixed in 0·1 % phosphotungstic acid solution, washed and dried. After drying, half of each membrane was scraped away by means of a fragment of razor blade so as to leave a uniform straight edge on one side of the membrane remaining on the cover-slip. A cover-slip bearing a piece of membrane and another clean cover-slip were silvered contemporaneously *in vacuo*, the thickness of the silver film being such as to give a reflexion coefficient of about 90 %. Pairs of cover-slips were now mounted with their silvered faces in contact in a suitable metal clamp

placed on the stage of a light microscope. This arrangement was now illuminated with a parallel beam of light from a mercury-arc lamp filtered through a monochromatic green filter and the resulting interference fringes observed with a low-power objective. From the displacement of the fringes where they cross the straight edge of the membrane the thickness of the dried membrane was estimated to be approximately 500 Å.

Taking the thickness of the membrane from interference measurements (which agree surprisingly well with estimates from shadow lengths) and taking figures for pore size and pore separation distance which are means between the measurements on supported and unsupported membranes, we are in a position to construct a sectional diagram of the nuclear membrane. Such a diagram is shown in figure 7.
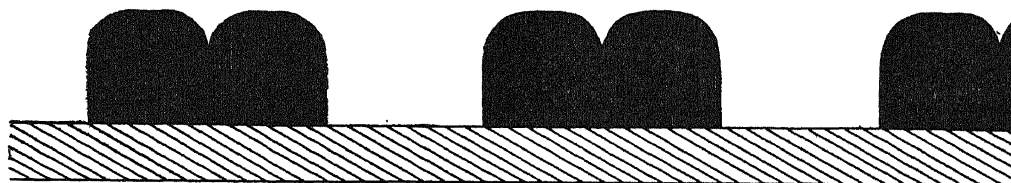
FIGURE 7. Reconstructed sectional diagram of the nuclear membrane, the section passing through the diameters of the pores. The a-layer is shown black, with indentations at the points of mechanical weakness where breakdown gives rise to 'annuli'. The b-layer is shown cross-hatched. The shape of the structural units of the a-layer is inferred only. The relative magnitudes of layer thickness, pore diameter and separation distance between pore centres have been directly determined. (Magn. × 500,000.)

4. DISCUSSION

There can be little doubt, from the evidence presented in the foregoing account, that the nuclear membrane of amphibian oocytes is a dual structure. The results of three different kinds of experiment agree in demonstrating that the two surfaces of the membrane are different in character. The external surface is porous; the evidence that this porous structure is not merely an artefact has been presented on p. 372. The internal surface gives no indication of heterogeneity of texture; it is the surface of a discrete layer which can exist in the absence of the external porous material.

At first sight it might be imagined that the thicker porous a-layer acts as a mechanical support for the thinner structureless b-layer. It seems probable that the b-layer determines the permeability properties of the nuclear membrane; the actual pores through which penetrating molecules pass the membrane must be well outside the limits of resolution of the electron microscope, since such molecules as egg and bovine plasma albumin, glycogen and gum acacia are unable to penetrate the membrane by ordinary osmotic means. It is conceivable that the absence of visible structure in the b-layer is itself a technical artefact, though this is very unlikely. It seems improbable that the visibly porous membrane can play a major part in determining the limits to permeability, though its existence may markedly reduce the surface area of the b-layer available for the passage of materials from nucleus to cytoplasm, or vice versa. In our manipulations of the nuclear membrane
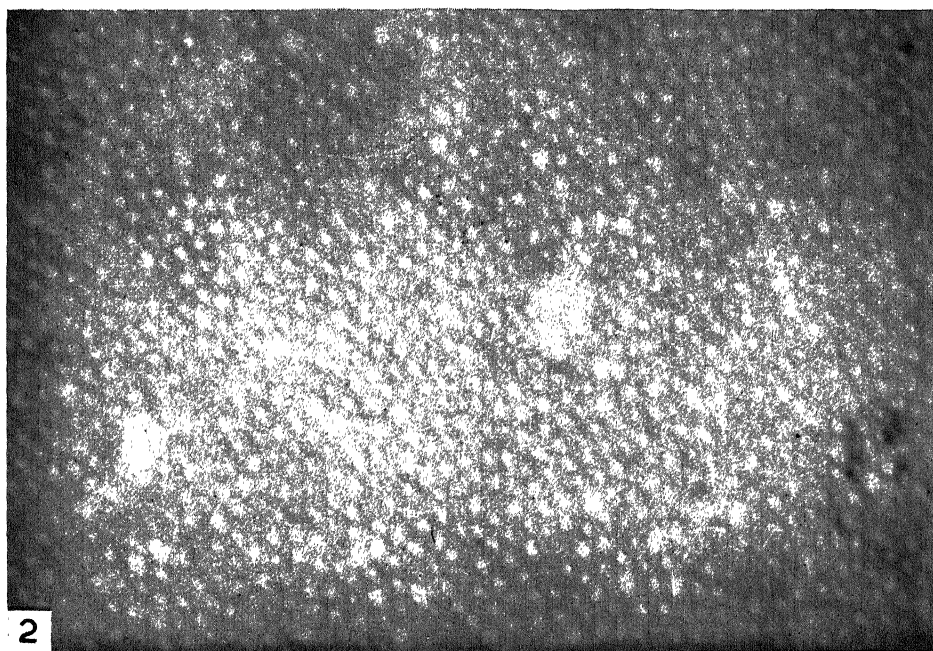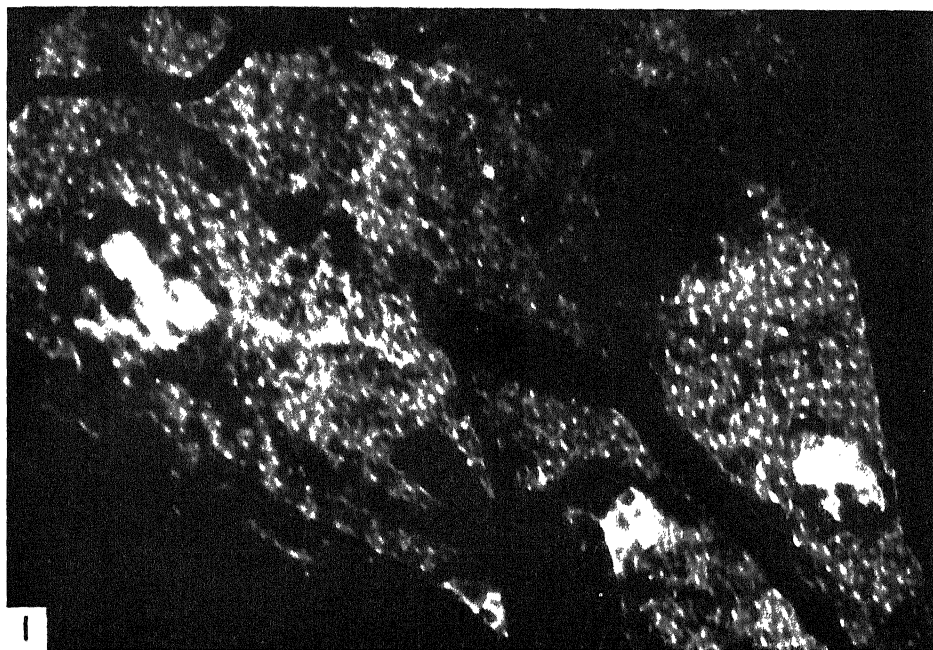
FIGURE 1. *Xenopus laevis*. Nuclear membrane isolated in distilled water, stretched over
copper grid and fixed for 15 min. in 2 % phosphotungstic acid. (Magn. × 26,000.)

FIGURE 2. *Xenopus laevis*. Nuclear membrane isolated in distilled water, mounted on Formvar
film with outer surface in contact with film, fixed for 2 min. in 0·1 % phosphotungstic
acid. (Magn. × 26,000.)
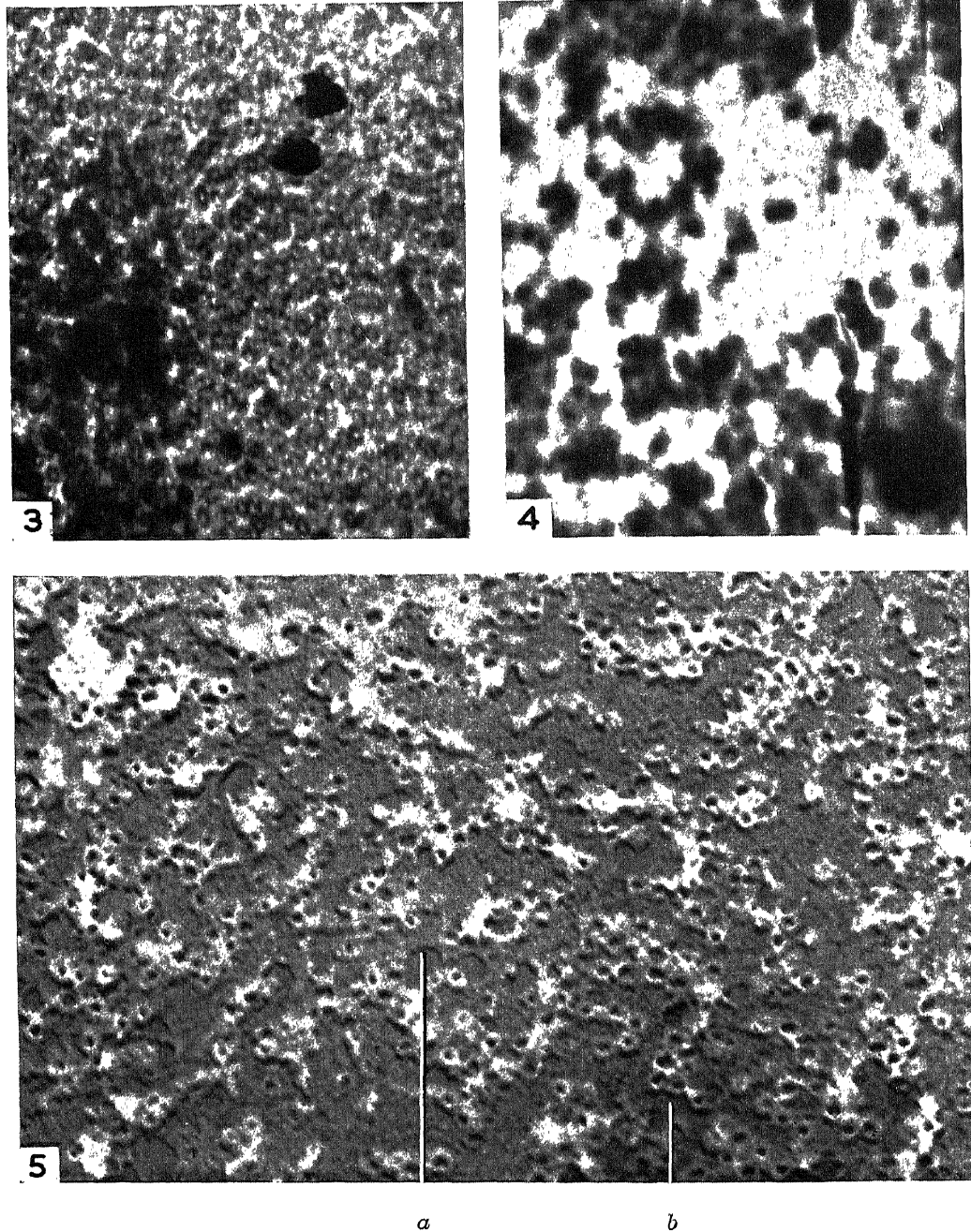
FIGURE 3. *Xenopus laevis.* Nuclear membrane isolated in 0·8 % sodium chloride, stretched over copper grid, washed in 10 % sodium chloride, fixed for 2 min. in osmium tetroxide vapour. (Magn. × 26,000.) Porous layer broken down to form annuli.

FIGURE 4. *Xenopus laevis.* Nuclear membrane isolated in 0·2 M-potassium hydrogen phosphate at pH 6·8, stretched on copper grid and fixed for 2 min. in 0·1 % phosphotungstic acid. (Magn. × 26,000.) Porous layer disrupted leaving debris supported by the continuous layer.

FIGURE 5. *Xenopus laevis.* Nuclear membrane isolated in distilled water, stretched on copper grid, fixed for 2 min. in 0·1 % phosphotungstic acid, backed on inner surface with Formvar film. Specimen inverted and shadowed through grid at an angle of 30° with palladium. Printed as a negative in order to show shadows dark. (Magn. × 26,000.) *a* indicates shadow of break in continuous layer on to Formvar; *b* indicates shadow of break in porous layer on to continuous layer.

we have not found that the presence of the *a*-layer markedly strengthens the membrane as a whole. The elastic properties and remarkable strength of the nuclear membrane seem to reside mainly in the *b*-layer, since these properties are almost as evident in nuclei isolated in phosphate as in those isolated in chloride solutions. It is true that membranes isolated in phosphate and stretched on grids collapse more readily on drying out than do those isolated in chloride, but this is hardly a fair test of the possible supporting function of the *a*-layer in life.

The *b*-layer, insoluble in distilled water, physiological salines and 10 % sodium chloride, is formed of protein (Callan, unpublished, from birefringence properties, qualitative tests and digestion experiments; see also Baud 1948, 1949*a*, *b*). The *a*-layer, too, consists at least in part of protein. We have evidence that it also contains a certain amount of lipoid. If unfixed membranes mounted direct on copper grids are exposed to fat extraction by alcohol and chloroform, the *a*-layer is quite normal in appearance. If, however, the membranes are mounted on stainless steel grids and then exposed to alcohol and chloroform, the characteristic appearance of the *a*-layer is lost and in its place we see the *b*-layer with a uniform covering of material derived from the *a*-layer, all trace of the porous structure having disappeared. This is quite striking evidence that lipoids are concerned in the molecular architecture of the *a*-layer, since it is known that copper ions in extremely low concentration (10^{-5}M) can inhibit lipoid extraction (Waugh & Schmitt 1940); this would explain the difference in behaviour of membranes on copper as opposed to stainless steel grids.

The presence of lipoid in the nuclear membrane was confirmed by the observation that it stains selectively with Sudan Black. Nuclei isolated in saline were transferred to a dilute solution of Sudan Black in 70 % alcohol and subsequently examined in water after rapid passage through 70 and 50 % alcohol. The stained regions are the membrane, and, incidentally, the borders of the nucleoli. It is, therefore, possible that the *a*-layer represents the 'couche lipidique périnucléaire' of Baud (1949*a*, *b*). Polarized light studies on the amphibian oocyte nuclear membrane provide no evidence for the presence of lipoids orientated with their hydrocarbon chains radial to the surface, nor, so far as we are aware, has this been demonstrated for any other kinds of nuclei (see, for example, Chinn 1938 and Schmidt 1936). Baud has been able to demonstrate positive birefringence with respect to the radius of the nucleus only in the cases of nuclei stained with chrysoidin; it may be that this treatment breaks down lipoprotein complexes, thereby allowing the lipoids to become arranged in a film at the surface of the nucleus, there producing the birefringence which Baud has observed. Even if, in life, the lipoids have a regular orientation with respect to the structural features of the *a*-layer, it is quite probable that this distribution would give rise to statistical isotropy of the membrane considered as a whole.

REFERENCES

Baud, C. A. 1948 *Bull. Histol. Tech. micr.* **25**, 112.
Baud, C. A. 1949*a* *Bull. Histol. Tech. micr.* **26**, 12.
Baud, C. A. 1949*b* *Bull. Histol. Tech. micr.* **26**, 99.
Bensley, R. R. 1938 *Anat. Rec.* **72**, 351.
Brachet, J. 1938 *Arch. exp. Zellforsch.* **22**, 541.
Brachet, J. 1940 *Arch. Biol., Paris*, **51**, 167.
Chinn, P. 1938 *J. Cell. Comp. Physiol.* **12**, 1.
Duryee, W. R. 1937 *Arch. exp. Zellforsch.* **19**, 171.
Duryee, W. R. 1938 *Biol. Bull. Woods Hole*, **75**, 345.
Duryee, W. R. 1941 '*Cytology, genetics and evolution*'. University of Pennsylvania Bicentennial Conference, Philadelphia.
Duryee, W. R. 1942 *Proc. 8th Amer. Sci. Congr.* **3**, 45.
Gersch, M. 1940 *Z. Zellforsch.* **30**, 483.
Gross, R. 1916 *Arch. Zellforsch.* **14**, 279.
Martin, A. & Tomlin, S. G. 1950 *Biochim. Biophys. Acta*, **5**, 154.
Pantin, C. F. A. 1946 *Notes on microscopical technique for zoologists*. Cambridge University Press.
Pischinger, A. 1937 *Z. Zellforsch.* **26**, 249.
Schmidt, W. J. 1936 *Protoplasma*, **32**, 193.
Tolansky, S. 1948 *Multiple beam interferometry*. Oxford University Press.
Waddington, C. H. 1938 *J. Exp. Biol.* **15**, 382.
Waugh, D. F. & Schmitt, F. O. 1940 *Cold Spr. Harb. Symp. Quant. Biol.* **8**, 233.
Williams, R. C. & Wyckoff, R. W. G. 1944 *J. Appl. Phys.* **15**, 712.

Osmoregulation in surviving slices from the kidneys of adult rats

By J. R. Robinson

Department of Experimental Medicine, University of Cambridge

The total amount of water in slices of the renal cortex of adult rats was determined after they had been exposed to saline solutions with concentrations from 0·58 to 0·03 os-mol./l.:

(*a*) When metabolism was suppressed by lowering the temperature to 0 to 4° C.

(*b*) When the slices were respiring in oxygen at 38·5° C.

(*c*) When respiration was inhibited by cyanide at 38·5° C.

(*d*) When metabolism was reduced by cooling to temperatures between 38 and 16° C.

The amount of water in slices at 0 to 4° C varied inversely with the concentration of the medium, and when this concentration was less than 0·58 os-mol./l. the slices contained more water than the tissue *in vivo*.

Slices respiring at 38·5° C in solutions more dilute than 0·58 os-mol./l. contained considerably less water than slices in the same solutions at 0 to 4° C, and dilution of the medium from 0·58 to 0·19 os-mol./l. produced a much smaller increase in the amount of water in the slices when they were respiring than when they were at 0 to 4° C.

The uptake of oxygen at 38·5° C was independent of the concentration of the medium between 0·45 and 0·12 os-mol./l.

Slices whose respiration was inhibited by cyanide in 'isotonic' (0·30 osM) solutions at 38·5° C contained more water than slices respiring freely in concentrations as low as 0·06 os-mol./l.

The changes produced by cyanide in oxygen uptake and in water content were both reversible.

When the uptake of oxygen was reduced by cooling to 30° C, the amount of water in the slices decreased.

Evidence is presented that all the changes in water content were almost complete in 2 min., and were then maintained for several hours.

These results suggest that respiration is more important than the osmotic pressure of the external medium in determining the amount of water in cells. They cannot be explained by the orthodox theory that mammalian cells are in osmotic equilibrium with their surroundings, and indeed they suggest that the osmotic pressure of the cell fluids is normally 50 to 100 % greater than that of the extracellular fluids.

All the observations can be explained if energy derived from respiration is used to expel water from the cells, so that a steady state is maintained in which the higher internal osmotic pressure causes water to diffuse into the cells as fast as it is pumped out. The energy required to maintain the observed amount of water in the respiring slices was calculated by a simple theoretical treatment. It was found to be proportional to the observed oxygen uptake and to be a small percentage of its energy equivalent.

INTRODUCTION

The higher animals have developed mechanisms for surrounding their cells with an extracellular fluid of fairly constant composition. Accordingly, when surviving tissues have been studied by manometric techniques, they have usually been suspended in solutions designed to imitate these extracellular fluids in composition and in osmolar concentration. Such solutions have been discussed by Krebs (1950). Their total concentration is approximately 0·30 os-mol./l., and there can be little doubt that this is similar to the concentration provided by the body for the cells in their normal environment. Opie (1949), however, found that considerably stronger solutions were required to prevent the swelling of a number of tissues removed from the body of the rat. Moreover, some of the cells lining the renal tubules must be able to function in contact with anisotonic solutions, for the urine is rarely isotonic with the plasma. Considerations of this kind led Siebeck (1912) to study the osmotic properties of the kidneys of frogs, and it is surprising that no comparable work was done until Conway, Fitzgerald & MacDougald (1946) reinvestigated the same organ.

It was therefore decided to investigate the behaviour of slices cut from the kidneys of normal adult rats, in media which preserved the cation ratios of the extracellular fluid but differed in osmolar concentration. The oxygen consumption and the amount of water in the tissue were measured under a variety of conditions. The amount of extracellular water in the slices was determined by means of inulin and sucrose, so that changes in the amount of intracellular water could be derived from the observed changes in total water. When it was found that the osmolar concentration of the medium made unexpectedly little difference to the oxygen consumption and to the water content of the respiring tissue, the effects of induced variations in oxygen consumption upon the amount of water in the slices were studied. The results could not be explained if the cells were assumed to be in static equilibrium with their surroundings, and an attempt has been made to explain the size of the cells under different conditions as the result of a steady state maintained by the continuous supply of energy. The amount of energy required has been calculated and related to the observed oxygen consumption.

Methods

(a) Media

All the media were modifications of the one (A_2) described by Robinson (1949) in which slices of adult rat kidney were found to maintain a steady uptake of oxygen for several hours. This solution had a total osmolar concentration of 0·30 and contained physiological amounts of sodium, potassium, calcium, magnesium, phosphate and glucose. Hypotonic solutions were made by diluting the glucose-free saline with distilled water. A hypertonic solution (0·58 osm) was made by mixing 0·308 m-sodium chloride (232 ml.), 0·308 m-potassium chloride (8 ml.), 0·110 m-calcium chloride (6 ml.) and 0·154 m-magnesium chloride (4 ml.), adjusting pH to 7·4, and adding 24 ml. of m/15-phosphate buffer of the same pH. This saline was diluted with distilled water to make another hypertonic solution containing 0·45 os-mol./l. The osmolar concentrations were calculated from the known constituents of the solutions without regard for activity coefficients by adding together the molar concentrations of all the ions present. Glucose, 100 mg./100 ml., was added immediately before use to all the solutions except some of those to be employed in experiments in which the metabolism of the slices was inhibited by chilling to 0 to 4° C. As its initial concentration was only 0·005 os-mol./l., and the amount consumed by the slices was not determined, glucose was left out of account in calculating the osmolar concentration of the solutions.

(b) Preparation of the slices

Rats were stunned by a blow on the head and bled from the carotid. The kidneys were removed at once, and slices approximately 0·3 mm. thick were cut from the cortex as described by Cohen (1945). In most cases the razor was moistened with the 0·30 osm saline, and the slices were washed in the same solution in a Petri dish. In order to determine the amount of water in the renal cortex in its natural state in the body, some slices were cut with a dry razor, blotted with hardened filter-paper (Whatman 541) to remove blood, urine and cell debris, and placed at once in tared drying tubes. A dry razor was also used to cut slices which were placed immediately in solutions of different concentrations in small bottles surrounded by broken ice. These slices were left in the chilled solutions for 15 to 30 min., with frequent shaking, and were then removed and blotted so that the water content of slices in which metabolism had been suppressed could be compared with that of slices which had been respiring in solutions of the same concentrations.

As the slices were approximately 0·3 mm. thick it is unlikely that more than 5 to 10 % of the cells were damaged by the razor. The damaged cells probably disintegrated in the course of the experiments as a result of shaking (Robinson 1949), so that although their presence must have introduced a systematic error into the determination of oxygen uptake, the slices taken for analysis at the conclusion of the experiments should have contained mainly undamaged cells.

(c) Determination of the amount of water in the tissue

At the end of the experiments the slices were spread out on filter-paper and gently blotted. They were then transferred with dry forceps to tared glass tubes and

weighed. The loss in weight after they had been dried overnight in an oven at 105° C was assumed to be the weight of total water in the tissue.

(d) Determination of the oxygen consumption

Barcroft manometers were used as described by Robinson (1949). The slices employed for this purpose were taken from the dish of saline into which they had been placed as they were cut, spread out on hardened filter-paper, and gently blotted. Groups of blotted slices, weighing 60 to 100 mg., were placed on small tared watch glasses, rapidly weighed, and placed in the manometer flasks. The respiratory rates were all calculated in $\mu l.O_2/hr./mg.$ of moist tissue initially placed in the flasks. At the end of the experiments the manometers were taken separately from the water-bath, and the slices removed as quickly as possible in order to avoid changes in their water content which might have occurred on cooling to room temperature.

In a number of experiments cyanide was used to inhibit the respiration of the slices, and its concentration was varied during the course of the experiments by one or other of the following devices. The concentration was reduced by allowing hydrogen cyanide to distil over gradually from the flasks into the potassium hydroxide in the centre cups. The concentration was raised by starting with a cyanide-free medium in the main flasks, and replacing the potassium hydroxide in the centre cups by a mixture of potassium cyanide and hydroxide which gave off hydrogen cyanide as it absorbed carbon dioxide.

(e) Short experiments

Determinations were made of the amounts of water in slices which had been immersed from 10 sec. to 5 min. in a number of solutions at different temperatures. The slices were lifted with forceps from the dish of 0·30 osm saline in which they had been prepared, and placed for periods measured with a stop-watch into 20 ml. of the experimental solutions in a small beaker separated by an air jacket from a larger one containing water heated by a microburner. Temperatures were measured by a thermometer dipping into the solutions. When these contained cyanide, they were stirred by hand. Otherwise a stream of oxygen was bubbled through them.

(f) Determination of the amount of extracellular water in the slices

Batches of 30 to 50 slices were shaken with 0·30 osm saline containing 1 % of inulin or sucrose. After time had been allowed for these to diffuse into the interstices of the tissue, the slices were removed from the solutions and blotted. Samples of four to eight blotted slices were dried to determine their water content. Similar samples were placed in tared centrifuge tubes, weighed, and extracted overnight with a measured volume of distilled water. Next day the proteins were precipitated with trichloroacetic acid, and the volume adjusted so that 2 ml. aliquots contained the inulin or sucrose extracted from 20 to 30 mg. of tissue. 2 ml. aliquots of 1 in 200 dilutions of the solutions in which the slices had been equilibrated were also taken for estimations of inulin or sucrose by the method of Bacon & Bell (1948) as modified by Cole (1949, unpublished). Blank determinations were made on slices which had been soaked in saline solutions which did not contain inulin or sucrose. The amount

of extracellular water was calculated as the space occupied by inulin or sucrose at the concentration in which either had been found in the external solution, and expressed in g./100 g. of moist tissue. Water in damaged cells penetrated by inulin or sucrose would be included in 'extracellular' water determined in this way. The maximum error if 10 % of the cells were damaged, and if all these cells remained attached to the slices and did not shrink through loss of their contents, would have been about 5 g./100 g. of tissue. In fact the error was probably much less than this, for the reasons stated above.

(g) Determination of the swelling of the tissue, and of changes in the volume of intracellular water

It was not practicable to recover slices quantitatively from the Barcroft flasks, blot, and transfer them to drying tubes. Swelling was therefore calculated from the observed percentages of total water by using the amount of water in the tissue *in vivo* as a standard. The only assumption involved is that when the slices swelled, they took up water without gaining or losing solids. Changes in the volume of intracellular water under different conditions were also derived from the total water content of the slices by making use of the values for extracellular water determined as just described. The method of calculation is given below in general terms, and formulae are derived for use later.

Let W_T be the normal percentage of total water in the tissue *in vivo*, and let W_T' be the percentage of water found in swollen slices. The weight of swollen tissue which contains the same amount of solid matter as 1 g. of normal tissue is therefore

$$S_T = \frac{100 - W_T}{100 - W_T'}\text{g.} \tag{1}$$

S_T will be referred to as the 'relative swelling' of the tissue. Hence, after swelling until it contained W_T' % of water, a portion of tissue which had weighed M_0 g. in the body would weigh $M_0\dfrac{100 - W_T}{100 - W_T'}$g., and contain $M_0\dfrac{W_T'}{100}\dfrac{100 - W_T}{100 - W_T'}$g. of water. Moreover, if W_E' % of the weight of the swollen tissue were made up of extracellular water, the portion of tissue under consideration would contain

$$M_0\frac{100 - W_T}{100 - W_T'}\frac{W_T' - W_E'}{100}\text{g. of intracellular water.}$$

Now if extracellular water had contributed W_E % of the weight of the tissue in the body, the original M_0 g. of tissue would have contained $M_0(W_T - W_E)/100$ g. of intracellular water. So that the relative volume of intracellular water in the swollen tissue (i.e. the volume of intracellular water in a quantity of tissue which *in vivo* had contained 1 ml. of intracellular water) would be

$$S_c = \frac{100 - W_T}{100 - W_T'}\frac{W_T' - W_E'}{W_T - W_E}\text{ml.} \tag{2}$$

The actual percentages of total water which were found under different conditions have been stated in presenting the results. The volumes of intracellular water

derived from these by means of equations (1) and (2) have been employed later in developing a theoretical treatment of the results.

RESULTS

(a) The amount of water in the renal cortex of the rat in vivo

Twenty-one groups of slices were found to contain 75·7 % (s.d. ± 0·9) of water.

(b) The amount of water in slices removed from solutions of different osmolar concentration at 0 to 4° C

Table 1, column 3, shows the average percentages of total water which were found in slices equilibrated at 0 to 4° C with solutions having concentrations from 0·03 to 0·58 os-mol./l. The actual osmolar concentrations are given in column 1, and the tonicities of the solutions, referred to 0·30 osm as isotonic (tonicity = 1·00), in column 2. As might have been anticipated, the amount of water in the slices was found to vary inversely with the concentration of the solutions in which they had been immersed. Moreover, slices from all but the most concentrated solution contained more water than the tissue in its normal state in the body. This observation confirms the work of Opie (1949), who had obtained similar results using solutions of sodium chloride instead of a balanced saline.

TABLE 1. SUMMARY OF RESULTS

con-centration (os-mol./l.)	tonicity	slices at 0 to 4° C total water (%)	s.d.	no.	slices at 38·5° C total water (%)	s.d.	no.	oxygen uptake (μl./hr./mg.)	s.d.	no.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0·03	0·10	88·7	0·8	8	87·5	0·75	5	2·8	0·3	5
0·06	0·20	87·7	0·9	9	83·5	1·15	10	3·6	0·55	10
0·12	0·40	85·0	0·8	8	81·9	1·0	10	4·2	0·5	10
0·19	0·63	83·7	0·8	10	78·5	0·7	9	4·2	0·3	9
0·30	1·00	81·6	1·3	43	77·5	0·8	44	4·2	0·4	44
0·45	1·50	78·4	1·5	11	75·4	0·5	7	4·5	0·3	8
0·58	1·92	75·8	1·8	15	76·5	0·8	7	3·5	0·4	7

(c) The amount of water in slices which had been respiring in solutions of different concentrations at 38·5° C

The average amounts of water found in the slices removed from the solutions in which they had been respiring in the Barcroft flasks are given in column 6 of table 1. The slices upon which these averages are based had respired for 2 hr. in solutions more concentrated than 0·12 osm, but they were removed from the three most dilute media after ½ to 1 hr. because their oxygen uptake was found to fall off if they were left in such dilute solutions for longer periods. A comparison of columns 6 and 3 of table 1 shows that slices that had been respiring at 38·5° C in all but the most concentrated solution contained less water than slices which had been shaken in solutions of the same concentrations at 0 to 4° C, although the difference was small when the concentration was 0·03 os-mol./l.

(d) *The oxygen uptake of slices respiring at* 38·5° C *in solutions*
of different concentrations

The uptakes of oxygen in μl./hr./mg. of moist tissue are shown in column 9 of table 1. The table includes only the results of experiments in which the gas phase was pure oxygen, and pH was controlled between 7·3 and 7·4 throughout. It will be seen that between 0·12 and 0·45 os-mol./l. the concentration of the medium had no significant effect upon the oxygen uptake. The effect of dilution from 0·30 to 0·19 osm was demonstrated particularly clearly by an experiment suggested by Dr G. D. Greville, in which the same slices were exposed to both concentrations. 1·7 ml. of the 0·30 osm medium was placed in the flasks of eighteen manometers. The side bulbs of nine of these contained 1 ml. of the same medium, but those of the other nine manometers contained 1 ml. of distilled water. After the uptake of oxygen had been recorded for an hour, the contents of the side bulbs were tipped into the main flasks, and readings continued for another hour. The results are given in table 2, and they show clearly that the addition of water after 1 hr. did not affect the oxygen uptake. Although the respiration of the slices was unaffected by varying the concentration of the medium from 0·12 to 0·45 os-mol./l., table 1 also shows that it was reduced by concentrations above and below these limits.

TABLE 2. EFFECT OF DILUTION ON RESPIRATION

conc. (os-mol./l.)	O_2 uptake in 1st hour			conc. (os-mol./l.)	O_2 uptake in 2nd hour		
	μl./mg.	s.d.	no.		μl./mg.	s.d.	no.
0·30	4·3	0·2	7	0·30	4·1	0·2	7
0·30	4·5	0·2	9	0·19	4·2	0·3	9

(e) *The relation between the oxygen uptake and the water content of slices*
respiring at 38·5° C *in isotonic solutions*

The results so far presented suggested that the amount of water which the slices took up from the different solutions depended upon their respiration. This point was further investigated by keeping the concentration of the medium constant at 0·30 os-mol./l. and varying the oxygen consumption. The kidneys of different rats provided slices with respiratory rates which varied between 3·4 and 5·0 μl./hr./mg. of tissue. Lower rates were obtained by using various concentrations of cyanide as an inhibitor, and also by filling the manometer flasks with air and with mixtures of hydrogen and oxygen instead of pure oxygen. The buffering of the medium proved to be inadequate in some of the latter experiments, and the pH fell towards 6. In a few additional experiments, therefore, the medium, with or without potassium cyanide, was maintained at pH 5·8 to 6·0 by a phosphate buffer, and the flasks were filled with oxygen.

The results of all these experiments are collected in figure 1, which shows how the amounts of water found in the slices at the end of the experiments were related to their rates of oxygen uptake. It is evident that the amount of water in the tissue varied inversely with its oxygen consumption. Moreover, when the percentages of
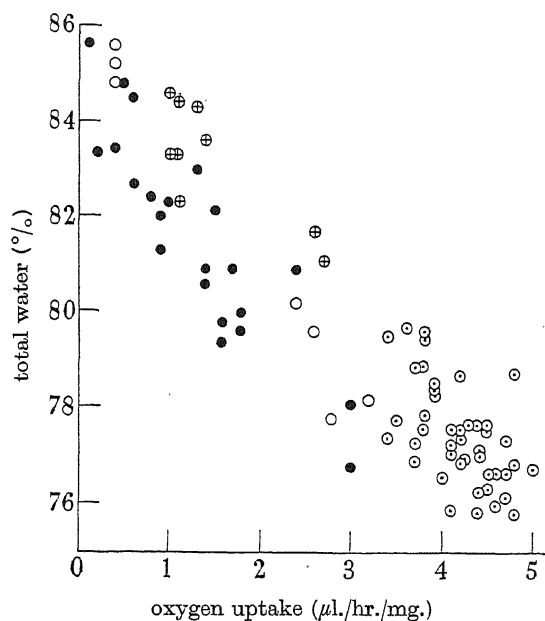
FIGURE 1. Water content and oxygen uptake of slices respiring in 0·30 osm solutions at 38·5° C. ⊙ respiration uninhibited. pH 7·3 to 7·4. Oxygen (2 hr.). ● in presence of KCN. pH 7·3 to 7·4. Oxygen (¼–2 hr.). ○ pH 5·8 to 6·0 ± KCN. Oxygen (2 hr.). ⊕ pH 6·0 to 7·0. Air or O₂/H₂ mixtures (2 hr.). (Duration of experiments in brackets.)
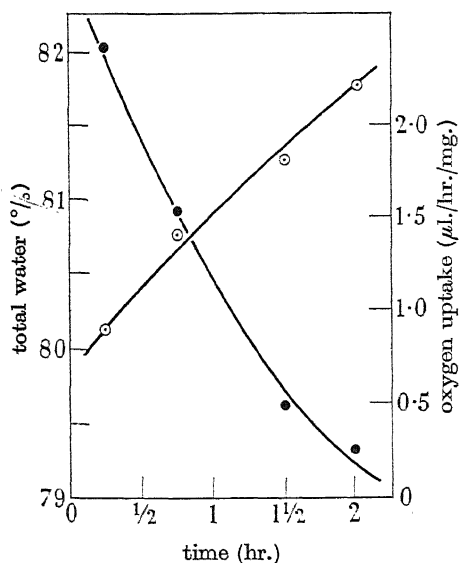
FIGURE 2. Water content and oxygen uptake of slices respiring in 0·30 osm solution with diminishing concentration of cyanide. ⊙ oxygen uptake. ● water content.
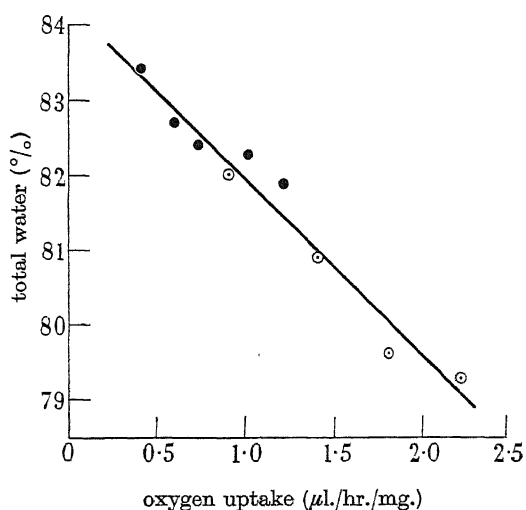
FIGURE 3. Water content and oxygen uptake of slices respiring in 0·30 osm solutions with rising ● and falling ⊙ concentrations of cyanide.

water plotted in figure 1 are compared with those listed in column 6 of table 1, it is clear that the slices took up more water from an 'isotonic' solution when their respiration was inhibited than they did from a solution of only one-fifth of this concentration in which they were respiring normally. How far these effects could be accounted for by metabolites produced in the absence of oxygen is discussed later (p. 392).

The increase which was found in the amount of water in the slices when their respiration was inhibited might have been caused by irreversible damage to the cells. Additional experiments were therefore carried out to test the reversibility of the effects produced by cyanide. Figure 2 shows how the oxygen uptake and the water content of the slices varied while cyanide was being gradually removed from a medium to which $M/200$-potassium cyanide had been added at the beginning of the experiment. The inhibition of the respiration proved to be reversible, and the slices evidently lost water as their oxygen uptake increased.

In figure 3 the amounts of water in these same slices, and also in those from another experiment during which the concentration of cyanide was increasing, have been plotted against the oxygen uptake. Figures 2 and 3 demonstrate that the effects of cyanide upon the respiration and upon the water content of the tissue were both reversible, and it seems reasonable to conclude from figure 1 that the amount of water in the slices was determined by their respiratory activity.

(f) The times taken to produce the changes in water content which were observed under different conditions

It took so long to measure oxygen uptakes in the Barcroft manometers that the experiments which provided the results in table 1 (column 6) and figures 1, 2 and 3 gave no indication of how long had been required to produce the changes in water content which were found at the end of the experiments.

Figure 4 shows the results of one of the 'short experiments' in which slices were taken from a dish of 0.30 osM saline in air at room temperature, and placed for from 10 to 240 sec. in a hypotonic (0.19 osM) solution at $38°$ C, (a) when respiration was uninhibited, and (b) in presence of $M/250$-cyanide. The final water content of these slices is plotted against the duration of their immersion in the warm solution. The average amount of water in slices removed from the dish at intervals during the experiment is plotted at zero time. Values from table 1 (columns 6 and 3) for respiring and chilled slices in 0.19 osM solutions are shown at A, off the time scale, for comparison.

Figure 5 shows the results of a precisely similar experiment in which the medium at $38°$ C was isotonic (0.30 osM). These experiments suggest that the changes found at the end of the longer experiments had been established during the first few minutes and then maintained for up to 2 hr. Figure 4 suggests further that when slices were placed suddenly into a hypotonic, oxygenated solution at body temperature, a decrease in their water content associated with the increased supply of oxygen occurred before water had time to diffuse in from the hypotonic solution. The rate at which water was taken up when the hypotonic solution contained cyanide

was presumably limited by the speed of diffusion of water through the intercellular spaces, and could not therefore be used to determine the permeability of the cell membranes to water.
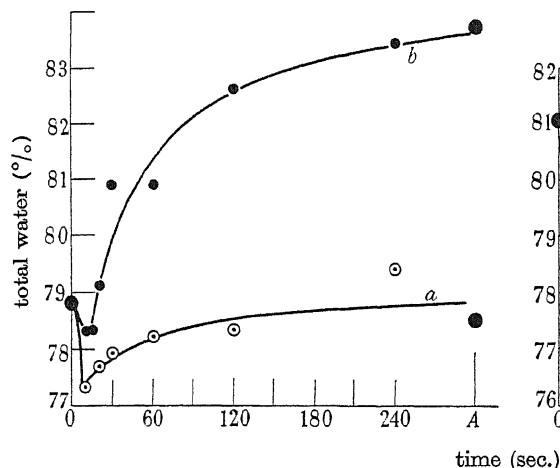
FIGURE 4. 0·19 OSM.   FIGURE 5. 0·30 OSM.

time (sec.)
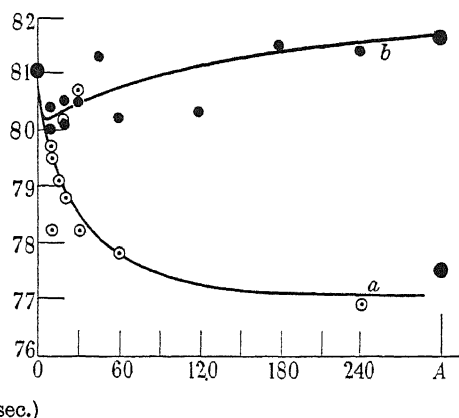
FIGURES 4 and 5. Time relations of changes in water content of slices in different solutions at 38° C. pH 7·4. ⊙ respiration uninhibited. ● in presence of M/250-cyanide. ● Points shown at *A* from table 1, columns 3 and 6. Average initial water content plotted at time = 0.

(g) The effect of temperature upon the oxygen uptake and the water content of slices respiring in isotonic solutions

Figure 6 shows how the oxygen uptake was affected by temperature. These results were obtained by varying the temperature of the water-bath, and recalculating the manometer constants for temperatures between 16 and 38° C. The dotted curve was calculated by assuming that the respiratory rate was 4·2 μl. O_2/hr./mg. of moist tissue at 38° C, and was reduced 2·5 times for each 10° fall in temperature (Q_{10} = 2·5). The amounts of water found in the slices for which figure 6 gives the oxygen uptake are shown by solid circles in figure 7. The remaining points in figure 7 are the results of short experiments in which the slices were placed for 5 min. at different temperatures in an oxygenated 0·30 OSM solution at pH 7·4. If the amount of water in the slices had depended on their oxygen uptake at other temperatures as it had been found to do at 38·5° C, the points in figure 7 should have been clustered around the dotted line (derived from figure 1) which passes above them. The slices evidently contained less water than might have been anticipated from their oxygen consumption.

This might have been because, at lower temperatures, respiration no longer played an important role in regulating the amount of water in the tissue. Ten groups of slices were therefore equilibrated at 18° C with an isotonic solution to which M/200-potassium cyanide had been added, and were found to contain 81·1 % (S.D. 0·4) of water. This is within the range given in table 1 for slices in 0·30 OSM solutions at 0 to 4° C, and higher than any value recorded in figure 7. A further

simple experiment also indicated that respiration was still controlling the amount of water in the slices at 18° C. A dozen slices from kidneys of the same rat were placed in each of two tubes containing 10 ml. of the same 0·30 osm solution at 18° C. Oxygen was passed through the contents of one tube, and hydrogen through those of the other for 5 min., after which the slices were quickly removed. Those taken from the oxygenated solution were found to contain 75·6 % of water, compared with 78·5 % found in the slices from the solution through which hydrogen was passed.
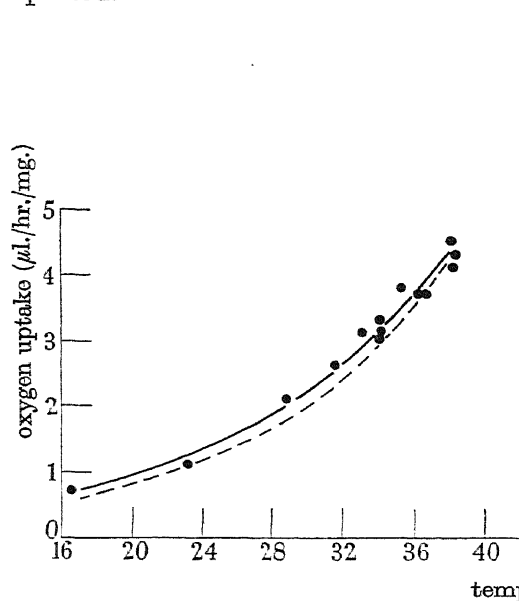
FIGURE 6. Effect of temperature on respiration of adult rat kidney slices in 0·30 osm solutions at pH 7·4. ● observed. − − − − calculated for 4·2 μl./hr./mg. at 38 °C, and $Q_{10} = 2·5$.
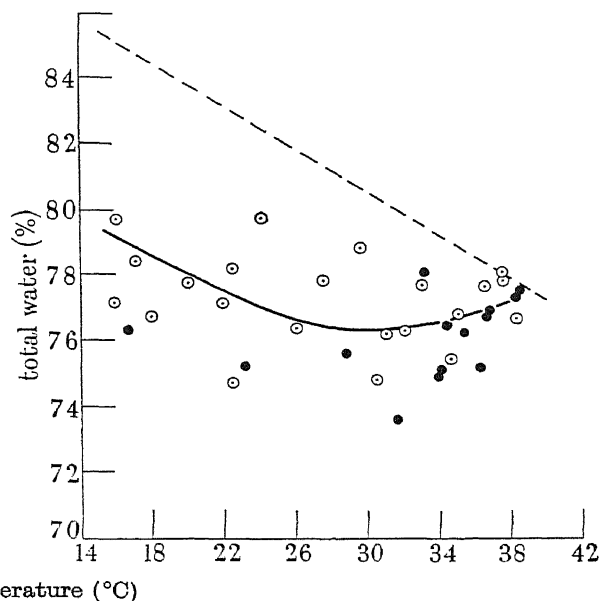
FIGURE 7. Effect of temperature on water content of rat kidney slices in 0·30 osm solutions at pH 7·4. ☉ immersed for 5 min. ● immersed for 1 to 3 hr. Dotted line shows trend expected if water content depended on respiration as in figure 1.

It therefore appeared that the percentage of water in the slices did still depend upon processes connected with respiration, even at these lower temperatures. Yet when the respiration was reduced by cooling, the amount of water in the tissue at first decreased, instead of increasing as it had done when respiration was reduced at body temperature. This discrepancy must remain unresolved until, at a late stage in the discussion, a theoretical treatment has been developed from which the curve drawn among the points in figure 7 was calculated.

(h) The amount of extracellular water in the slices

Preliminary experiments illustrated in figure 8 showed that the inulin space of slices placed in 0·30 osm saline at 0 to 4° C increased rapidly for the first 30 min., and then very slowly during the next 3 days. Consequently in the experiments to determine the amount of extracellular water in the slices, they were allowed 2 to 3 hr. to come into equilibrium with solutions containing inulin. In twenty-four

such experiments the inulin space was found to average 26·8 g./100 g. of moist tissue, which contained 82 % (s.d. 1·6) of water. Although both the inulin space and the total water had varied considerably in different experiments, their variations did not appear to be correlated, which suggested that the extracellular water might make up a fairly constant fraction of the weight of the slices irrespective of their degree of swelling. Two further experiments were made to test this hypothesis. First, sucrose was allowed to permeate slices for periods from ½ to 1 hr. at room temperature. Twelve groups of slices which contained on the average 77·5 % of water showed a sucrose space of 26·4 % (s.d. 0·8). In a second experiment, 1 % inulin was added to the 0·30 osm medium in which slices were respiring in the Barcroft manometers, glucose being omitted from the medium to reduce the blanks in the estimation of inulin. Six sets of slices which contained on the average 75·4 % of water showed an inulin space of 25·3 (s.d. 1·5) g/100 g. of tissue.
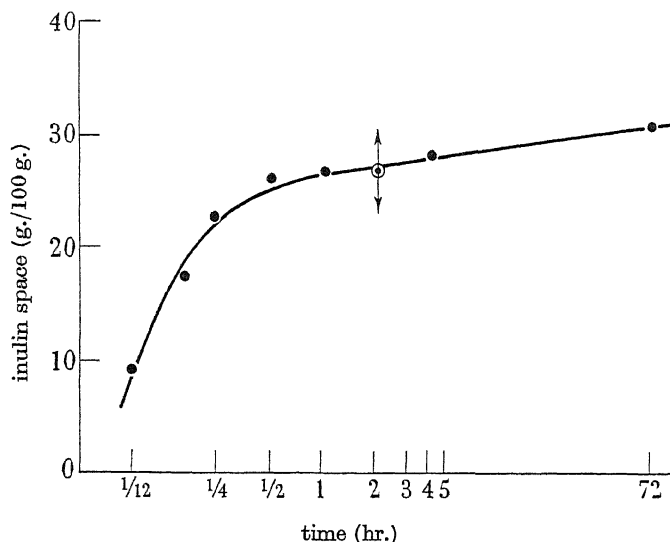
FIGURE 8. Entry of inulin into rat kidney slices in 0·30 osm saline at 0 to 4° C. ● individual observation. ⦶ average of 24 experiments (arrows show ± s.d.). Logarithmic time scale.

These results taken together suggest that about 26 % of the weight of the rat's renal cortex was made up of extracellular water, so that about 35 % of the water in the tissue *in vivo* was extracellular. This is comparable to the figure of 41 % which was found by Conway & Fitzgerald (1942) in the kidney of the rabbit. Since extracellular water was a constant fraction of the weight of the slices, the interstices must have expanded in the same proportion as the tissue as a whole when the slices swelled. This is not unreasonable, for the slices were thin, had no restraining capsule, and were seen to increase in all dimensions when placed in very dilute solutions.

(*i*) *The calculated changes in the weight of the slices, and in the amount of intracellular water*

All the symbols except W_T' (the percentage of total water in the slices) in equations (1) and (2) on p. 382 may now be replaced by experimentally determined constants.

If W_T, the percentage of water in the kidney *in vivo*, is taken to be 76 (p. 383), equation (1) gives the relative swelling of the tissue as a whole as

$$S_T = \frac{100-76}{100-W'_T} = \frac{24}{100-W'_T}.$$ (3)

Further, as extracellular fluid was found to make up 26 % of the weight of the tissue under all conditions of swelling,

$$W_E = W'_E = 26,$$

so that 100 g. of renal cortical tissue in its normal state contained 50 g. of intracellular water, and equation (2) gives the relative volume of the intracellular water as

$$S_c = \frac{100-76}{100-W'_T}\frac{W'_T-26}{76-26} = \frac{24}{50}\frac{W'_T-26}{100-W'_T}.$$ (4)

Values of S_T and S_c calculated from (3) and (4) are plotted against the percentage of total water W'_T in figure 9, which shows that the cell water increased more in proportion to its original amount during swelling than did the tissue as a whole.
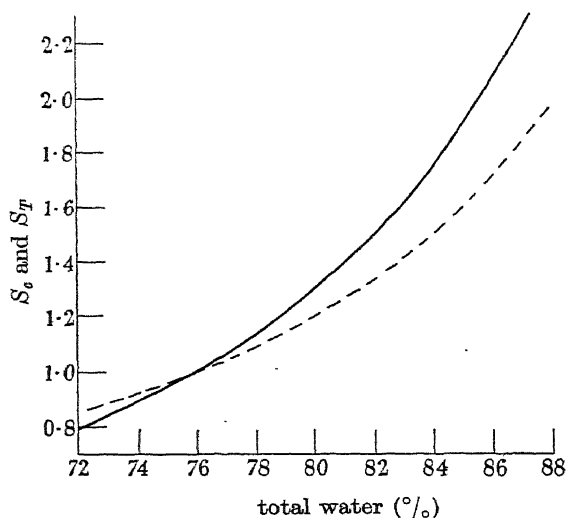
FIGURE 9. Relative swelling of whole tissue S_T (------), and relative volume of cell water S_c (————), corresponding to different percentages of total water. Calculated from equations (3) and (4).
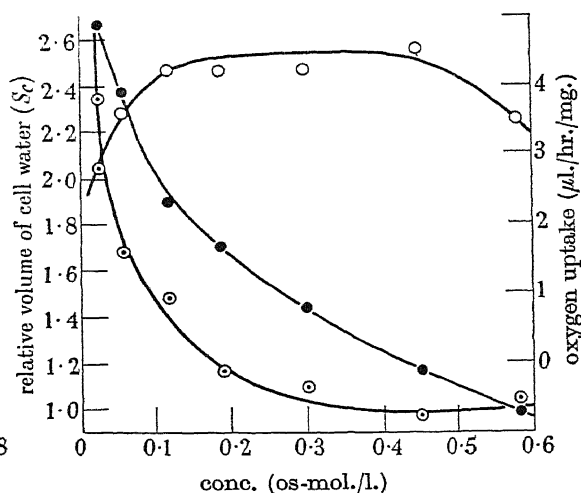
FIGURE 10. Effect of concentration of medium on: (1) volume of intracellular water in slices at 0 to 4° C ●; (2) volume of intracellular water in slices at 38·5° C ☉; (3) oxygen uptake of slices at 38·5° C ○.

The reversibility of the changes in water content which have been described also suggests that they were due mainly to changes in the intracellular water of the tissue. Opie (1949) was led to a similar conclusion by histological examination of the swollen slices he had used.

Equation (4) was used to calculate the relative volumes of cell water corresponding to the percentages of total water in columns 3 and 6 of table 1. The results are shown plotted against the concentrations of the solutions in figure 10, and the respiratory

rates from column 9 of table 1 are shown on the same diagram, which thus provides a summary and recapitulation of the data in the table. The effect of respiration is clearly demonstrated by the separation of the swelling curves for the chilled and the respiring slices. Moreover, these curves came together when extremes of concentration at either end of the range interfered in some way with the metabolism of the slices.

The product of the relative volume of intracellular water and the concentration of the medium was not constant, and the results in figure 10 could only be made to fit an equation of the type $P(V - b) = $ constant, by a correction for 'non-solvent volume' of the order reported by many workers on a variety of tissues. It appears from the discussion of Brooks & Brooks (1941) that the significance of such a correction is so uncertain that this aspect of the results has not been further considered.

DISCUSSION

(a) General discussion of the results and their significance

The alterations which have been described in the amount of water in slices of adult rat kidney were so large that they must have reflected changes in most if not all members of the mixed population of cells in the slices. Preliminary experiments have suggested that respiration affected the amount of water in the cells of the liver and diaphragm of the adult rat, in the kidney cells of the newborn rat and of the adult dog in a qualitatively similar manner. Moreover, Stern, Eggleston, Hems & Krebs (1949) found that slices of the kidney, liver, brain, lung and spleen of guinea-pigs were heavier under anaerobic conditions than when they were respiring in oxygen. In the case of kidney slices the difference was 52 %, in close agreement with the present observations on rat kidney. Another incidental observation of the same kind was made by Conway & Fitzgerald (1942), who studied the permeation of slices of rabbit kidney by inulin, and found changes in the water content of the tissues in the presence of cyanide, which were similar to those reported above for the rat. Respiration therefore appears to play an important part in determining the amount of water in a number of mammalian cells. (*Note.* That such behaviour is not confined to mammalian cells is suggested by work on organisms with contractile vacuoles (Kitching 1938, 1948) which was not seen until after this paper had been written.)

Exchanges of body water between the intracellular and extracellular compartments have usually been interpreted on a view put forward by Peters (1935). This view was most concisely stated (Peters 1944) in the words: 'Despite the extreme differences in composition of the contents of its various compartments, a uniform osmotic pressure prevails throughout the fluids of the body. It follows that the membranes between these compartments must universally permit the free passage of water, and that exchanges of water among the compartments must be determined by the hydrostatic and osmotic forces within them.' If, however, many cells in the body behave as some have now been found to do *in vitro*, the metabolism of the tissues may be a factor of even greater importance than these physical forces. From the results which have just been presented, shifts of water would be expected

to be determined by the osmotic pressure of the extracellular fluids *only if the metabolism of the cells remained constant*. In other words, the water balance within the body under physiological conditions should depend only partly upon the concentration of the extracellular fluids. Physiologists and clinicians should not be surprised to find paradoxical alterations of fluid equilibrium in the presence of metabolic disturbances. Moreover, a link between the volume of a cell and its metabolism may shed light on the mechanism of some of the mysterious swellings and degenerations encountered by pathologists.

The trend of the points in figure 1 which were derived from experiments at pH 7·4 shows that the slices would have contained about 84 % of water if their respiration had been completely inhibited. Hence from equation (4) or figure 9 the volume of the cell water had increased to 1·74 times its initial value. To account for this swelling in an isotonic solution it would be necessary, on the theory that the cells were in osmotic equilibrium with their surroundings, to suppose that the amount of osmotically active material inside them had increased to a similar extent. The change from aerobic to anaerobic metabolism might be expected to lead to an alteration in the number of molecules of organic metabolites in the cells, but the magnitude of this alteration cannot even be guessed, and Peters seems to have regarded the total base of the cells as of far greater importance than their organic constituents in determining the movement of water. Moreover, to explain the relatively small changes which have now been found in the volume of the cells respiring in solutions ranging from 0·58 to 0·19 os-mol./l. (table 1 and figure 10), it would be necessary to suppose that the cells could, in a few seconds (figure 4), vary the amount of osmotically active material inside them almost in proportion to the dilution of the medium. It is not easy to imagine how this could be done unless the membranes allowed free passage to solutes, in which case the volume of the cells could hardly be determined by the osmotic pressure of their surroundings. The orthodox theory seems here to have broken down just where it had usually been most successful. For its greatest value lay in predicting movements of water about the body in response to changes in the osmotic pressure of the extracellular fluids.

The theory has sometimes failed to explain shifts of water which were found to accompany changes in the concentration of intracellular base. Thus Gaudino & Levitt (1949) described how they administered desoxycorticosterone acetate to dogs and found a large increase in the concentration of sodium and potassium in the cells, whilst at the same time about one-third of the water moved out into the extracellular compartment. They commented that in order to preserve osmotic equilibrium there must have been changes in osmotically active cell constituents other than sodium and potassium and went on: 'Alternatively, if osmotic equilibrium between the cells and the extracellular fluid does not obtain, it would seem necessary to suppose that tissue cells generally are capable of conditioning the distribution of water independently of osmotic pressure, a rather implausible assumption.' But Adolph (1947) closed a review on water metabolism with the words: 'The forces that guide transfer may be more numerous than present theories of osmotic and hydrostatic pressure suggest'. The assumption that Gaudino & Levitt rejected is not implausible in the light of the results of the present work.

These lend strong support to the conclusion of Stern *et al.* (1949) that respiration was an important factor in regulating the uptake of water by the slices; and that the dominant factor in fluid exchange is not some physical property of a membrane, but a 'mechanism dependent on the supply of energy.'

Peters (1937–8) himself recognized that there were discrepancies between the concentrations of total base in the serum and in the cells, and in order to reconcile these discrepancies with the theory he postulated that by a sort of 'osmotic buffer mechanism' a variable part of the base inside the cells was bound in some osmotically inactive form. With this additional assumption the theory lost the attractive simplicity which had been one of its greatest merits. Moreover, so long as the actual amount of base which must be considered osmotically inactive cannot be determined by independent methods, the statement that the cells have the same osmotic pressure as their surroundings must remain an assumption, and, indeed, an assumption which could neither be proved nor disproved even if it became possible to make a complete chemical analysis of the contents of the cells. In fact, workers who have tried, like Sabbatani (1901), to measure the osmotic pressures of the cells of parenchymatous organs directly, have not found them to be the same as those of the extracellular fluids, but some 50 to 100 % greater. Thus Gömöri & Molnár (1932) found that the freezing-points of a number of organs of the rabbit were considerably lower than that of the serum, but that they approached the latter closely in the terminal stages of water intoxication. Later, Gömöri & Frenreisz (1937) found that the difference in freezing-points between the organs and the serum of cats became greater when dehydration was produced by experimental pyloric stenosis. The variations which were observed in these experiments suggest that the hypertonicity of the cells which was found by the use of Sabbatani's method was not an artefact due to substances of low molecular weight produced by autolytic processes in the dead tissues.

The cells of the tissue slices used in the present work were most likely to have been in osmotic equilibrium with their surroundings at 0 to 4° C when their metabolism was inhibited by chilling, and at 38·5° C when the supply of energy from respiration was cut off by poisoning with cyanide. In the former case they could only be prevented from swelling by placing them in solutions of 0·55 to 0·60 osm concentration (figure 10), and this should therefore be the normal osmolar concentration of the cells. When they were placed in 0·30 osm solutions at 0 to 4° C, the slices swelled until they contained 81·6 % instead of the normal 76 % of water, so that from figure 9 the relative volume of intracellular water was 1·45. If the cells in their normal state had contained the same amount of solute dissolved in 1·45 times less water, the normal concentration of the intracellular fluid would have been $1·45 \times 0·30 = 0·44$ os-mol./l. Similarly, since for slices poisoned with cyanide in a 0·30 osm solution at 38·5° C, the relative volume of cell water, S_c, was 1·74, the normal concentration of the intracellular fluid should have been $1·74 \times 0·30 = 0·52$ os-mol./l., which lies between the two estimates based on the amount of water in chilled slices. The discrepancies between these estimates may have been caused partly by the effect of temperature upon the hydration of colloids, and partly by binding of some of the water in the cells in a form which did not allow it to move in response to

changes in external osmotic pressure. Nothing is known about the magnitude of the first of these effects, and the second one has been discussed by Hill (1931). Different methods of measurement have led to quite different results, for the osmotic properties of many cells have suggested that they contained considerable amounts of 'bound' water, whereas Hill's (1930 b, c) determinations of vapour pressure suggested that most of the water in frog muscle was 'free'. The possibility that some of the water in the kidney slices was bound has, therefore, been left out of account in the theoretical treatment of the results. All that can be said at present is that the three estimates which have just been made, like the results of workers with cryoscopic methods, indicate that the osmotic pressure of some cells in the body is 50 to 100 % greater than that of the extracellular fluids with which on the orthodox theory they should be in osmotic equilibrium.

Such a hypertonicity of respiring cells could not exist as a true equilibrium, but it could be maintained as a 'dynamic equilibrium' or 'steady state' by the continuous supply of energy. Hill (1930 a, 1931) was among the first to suggest that many apparent 'equilibria' in biology are really steady states. True equilibria, studied by classical thermodynamic methods in 'closed' systems isolated from their surroundings, persist indefinitely and require no energy to maintain them. But living organisms, as pointed out by Bertalanffy (1950), are 'open' systems which continually exchange matter as well as energy with their surroundings. True equilibria are not often found in such systems, although steady states occur which resemble equilibria in that they do not change with time, but differ from them in that they can only be maintained by a continual supply of energy. If energy ceases to be available, these false 'equilibria' change with time until they reach states of true equilibrium in the classical sense. Hardly anything is known about the detailed theory of the 'steady state', but Bertalanffy (1950) has reviewed attempts to apply it to the study of the accumulation of ions in cells, of growth, and of cell division. The remainder of this paper contains an attempt to calculate the amount of energy required to maintain the water balance of the slices as a steady state.

(b) Theoretical interpretation of the results

(1) Introduction

On the assumption that when they were respiring, the cells were *not* in osmotic equilibrium with their surroundings, water should have been diffusing into them at a rate proportional to the difference in osmolar concentration between the intracellular fluid and the medium. This difference in concentration could then be kept up by a mechanism which made use of energy derived from the resting metabolism of the cells to expel water from them. Such a process could account for the 'osmotic buffering' of respiring cells (figure 10) dynamically, without the assumption that any intracellular solute was bound in an osmotically inactive form. The cells would only be expected to swell in hypotonic solutions if the rate at which water was expelled could not be increased in proportion to the concentration gradient causing it to diffuse into them. The results illustrated in figure 1 could be explained by supposing that when the respiration was depressed by cyanide, less energy was available for this purpose, so that the outward transport of water became slower.

The cells therefore swelled, and by so doing reduced the concentration gradient which was driving water into them, until a point was reached when the rate at which water diffused in had fallen to that at which it was being pumped out. When respiration had completely stopped and no energy was available, the outward transport of water ceased, and the volume of the cells was determined by a true osmotic equilibrium.

As an introduction to the calculation of the amount of energy which would be required for this transport of water, the assumptions which have been made will be discussed in connexion with certain alternative conceptions of the detailed mechanism of the process.

First, it has been assumed that the permeability of the membranes to water did not change except when the temperature was varied, and that they remained semi-permeable at all temperatures. Blumenthal (1927) explained an increase in the rate of swelling of *Arbacia* eggs in diluted sea water in the presence of hydrogen cyanide by an increase in permeability. Such a change might have been invoked to explain the results in figure 1, because although the swelling occurred in an 'isotonic' solution, the interior of the cells in their normal state was assumed to be hypertonic. However, it is simpler to interpret the results in terms of the known effect of cyanide on respiration. Incidentally, if the membranes had been seriously damaged by cyanide, solutes might have been expected to leak out of the cells, and they should have shrunk instead of swelling. It has also been assumed that only water moved in and out of the cells when their volume changed. In fact, it is not justifiable to assume that the membranes were absolutely impermeable to solutes, for the differences in composition between intracellular and extracellular fluids were probably also maintained as steady states by the expenditure of energy, and not by the impermeability of the membranes. Provided that these differences in composition were maintained, however, their effect upon the osmotic movement of water would probably be much the same as if the membranes were truly semi-permeable. This point is linked up with a third assumption, that the active process was primarily a transfer of water. A. L. Hodgkin (private communication) suggested that the changes might have been due to movements of water accompanying an active transport of sodium by a process connected with respiration. Conway *et al.* (1946), moreover, found evidence that sodium was actively extruded from the cells of the distal tubules in the frog's kidney. This process could be stopped by cyanide, and the water which then entered the cells along with sodium produced a considerable increase in the weight of the tissue. The changes took place much more slowly than those now being discussed in slices of rat kidney, which may suggest that the latter were due mainly at any rate to the movement of water. To test this, a number of measurements were made of the amount of water in slices which had been in bottles at 0 to 4° C and in the Barcroft flasks at 38·5° in solutions from 0·58 to 0·12 os-mol./l., in the prepartion of which choline chloride had been substituted for sodium chloride. Slices were also studied in a 0·30 OSM sodium-free solution in the presence of a variable concentration of potassium cyanide. The amounts of water in these slices were somewhat less than the averages from the experiments which have been presented in detail, but were within their range of

variation. The rates of oxygen uptake found for these slices were also within the normal range. It is interesting to note that slices respiring in 0·30 osm chloride-free media (Robinson 1950) also had normal respiratory rates and contained the same amount of water as slices respiring in the 0·30 osm solutions used in this work. These results suggest that though the active transport of sodium or other ions may have played some part in the movement of water through the cell membranes, water was also transported as such, and it should be justifiable as a first approximation to calculate the energy required to move water alone. No assumption has therefore been made about the way in which water was extruded from the cells.

Lastly, the mol fraction of water has been used in place of its activity in the formulae employed to calculate the rate of diffusion of water into the cells and the amount of osmotic work that would have to be done to expel it. As all solutions were more dilute than 0·60 osm, the difference between molal and molar concentrations was neglected (Glasstone 1947), and the mol fraction of water in a solution of osmolar concentration m has been taken to be $= 1 - m/55$ (1 kg. of water contains 55 moles approx.).

(2) *Calculation of the amount of osmotic work required to maintain the steady state*

According to Davson & Danielli (1943, p. 40) the rate at which an uncharged substance penetrates into a cell may be obtained from the simple formula

$$\frac{dS}{dt} = kA(C_o - C_i),\tag{5}$$

which is based on the diffusion equation of Fick. A is the surface area of the cell, k is a permeability constant which includes the thickness as well as the absolute permeability of the membrane, C_o and C_i are the external and internal concentrations of the substance to which the permeability constant applies, and dS/dt is the rate of penetration, which may be obtained in mol./hr. by a suitable choice of the units in which k and A are expressed. From (5), the rate of penetration of water into a cell maintaining an internal osmolar concentration m_i while respiring in a medium of concentration m_o is

$$kA\left[\left(1 - \frac{m_o}{55}\right) - \left(1 - \frac{m_i}{55}\right)\right] = \frac{kA}{55}(m_i - m_o).\tag{6}$$

Further, according to Borsook & Winegarden (1931), the amount of energy required to transfer n mols of water reversibly from a solution in which its mol fraction was N_1 to one where the mol fraction of water is N_2 is given by

$$nRT\ln\frac{N_2}{N_1},\tag{7}$$

where R is the gas constant and T the absolute temperature. Hence the amount of energy required to expel n mols of water from the cell is

$$nRT\ln\frac{1 - m_o/55}{1 - m_i/55}.\tag{8}$$

Since $m_i/55$ and $m_o/55$ are small compared with unity, the amount of osmotic work required to expel n mols of water is

$$W = \frac{nRT}{55}(m_i - m_o). \tag{9}$$

Moreover, to keep the cell in the steady state in which the osmolar concentration inside it is m_i, the actual amount of water expelled in a given period of time must be the same as the amount diffusing in, and this is given by (6). With the substitution of the expression (6) for n, equation (9) becomes

$$W_s = \frac{kART}{3025}(m_i - m_o)^2, \tag{10}$$

where W_s is the rate at which osmotic work would have to be done to maintain a concentration of m_i os-mol./l. in a cell surrounded by a solution of osmolar concentration m_o. The relative amounts of osmotic work which would have been required to maintain the water content of the cells which had been observed under different conditions were then calculated as follows: The relative volume of cell water, S_c, corresponding to the observed percentage of water in the slices was found from figure 9. Then, since the normal concentration of the intracellular fluid at body temperature had been estimated to be 0·52 os-mol./l. (p. 393), m_i was set equal to $0·52/S_c$. The average area of the cell surface under normal conditions was unknown but could be assumed to be constant, and as an arbitrary correction for the effect of swelling, A in equation (10) was replaced by $A'S_c^{\frac{2}{3}}$. This gave equation (10) in the modified form

$$W_s = \frac{kA'RT}{3025}S_c^{\frac{2}{3}}\left(\frac{0·52}{S_c} - m_o\right)^2, \tag{11}$$

which besides constants contained only m_o, the concentration of the medium used in any experiment, and S_c, which could be found from the observed percentage of total water in the slices. The values of the expression

$$S_c^{\frac{2}{3}}\left(\frac{0·52}{S_c} - m_o\right)^2 = \text{'relative osmotic work'} \tag{12}$$

were then calculated for percentages of total water from 73 to 88 and for the concentrations of all the media which had been used, and plotted against the amount of water in the tissue (figure 11). The amount of osmotic work required to keep the cells in a steady state in which the tissue contained any observed percentage of water was then determined by reference to the curve for the appropriate concentration of the medium in figure 11. The results for all slices respiring in 0·30 osM solutions in oxygen at pH 7·3 to 7·4 (cf. figure 1) have been plotted in figure 12 against the respiratory rates of the same slices. Despite a considerable scatter, it is clear that the amount of osmotic work required to maintain the observed water content of the slices was approximately proportional to that part of their respiration which was sensitive to cyanide. Or, in other words, that the amount of water actually found in the slices could have been maintained by the expenditure of a quantity of energy proportional to their respiration.

The amounts of osmotic work necessary to maintain the observed water content of slices respiring in anisotonic solutions have been compared in figure 13 with the amount of osmotic work which from (12) would have been needed to maintain in all concentrations the water content of 77·5 % ($S_c = 1·10$ from figure 9) which was found in slices respiring in isotonic (0·30)-osm solutions. The respiratory rates have been shown in the same diagram. It will be seen that moderate hypotonicity of the medium was accompanied by an increase in osmotic work. Although the actual oxygen uptake did not increase, the slices in the hypotonic solutions were consuming more oxygen than slices of the same water content respiring in isotonic solutions (figure 1). The slices may have swelled somewhat in the hypotonic solutions, not because they were in osmotic equilibrium, but because the increase in osmotic work was not sufficient to prevent swelling altogether. At great dilutions, when respiration was impaired, the osmotic work indicated by the amount of water found in the slices fell off again, just as in figure 10 the difference in swelling between the chilled and the respiring slices had diminished.
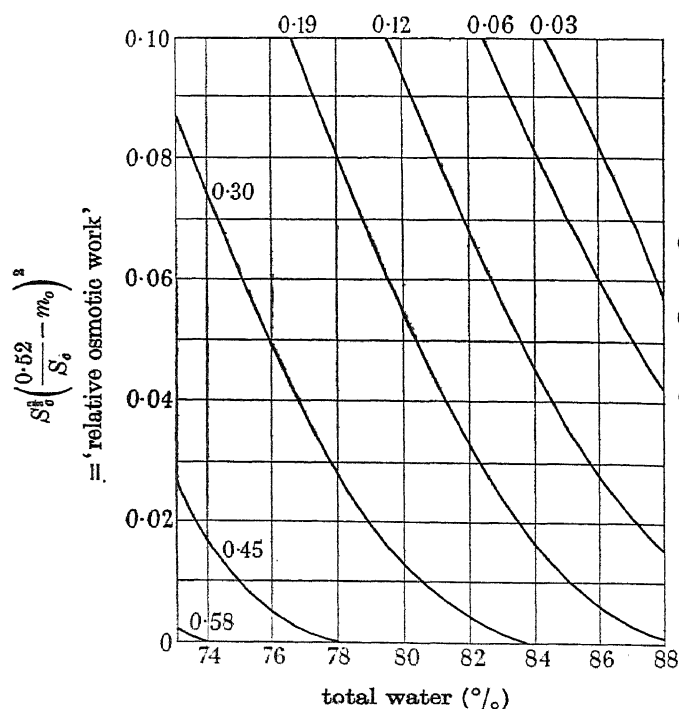
FIGURE 11. 'Relative osmotic work' required to maintain water content of slices as a steady state in solutions of different concentrations (m_o).
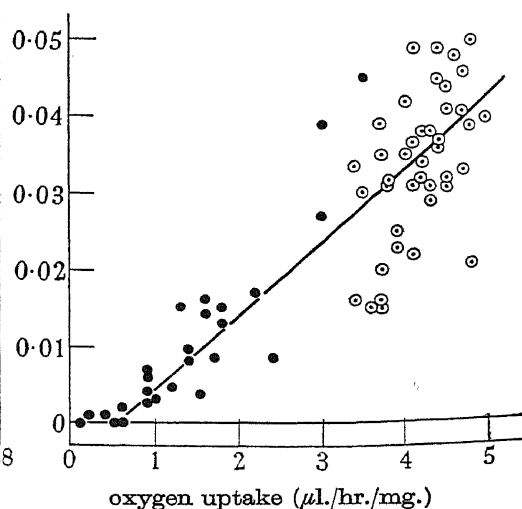
FIGURE 12. 'Relative osmotic work' required to maintain observed water content of slices respiring in 0·30 osm solutions at 38·5° C and pH 7·4. ⊙ with no inhibitor. ◕ in presence of cyanide.

(3) The effect of temperature on the amount of water in the slices

The effects of temperature upon dynamic phenomena are sometimes unexpected, and may be used to provide crucial experiments to test a theory (Robinson 1939). The fact that the slices did not take up water when their respiration was inhibited

by cooling to 30° C was in striking contrast with the changes which had occurred when other methods had been employed to depress respiration, but it can be predicted from the simple theory which has been developed. The absolute temperature appeared explicitly in equation (9), but in such a manner that the effect of variations between 16 and 38° C would be unimportant. Temperature, however, appeared in a masked form in equation (10), in the permeability constant k. It is a peculiarity of living membranes that their permeability to water has a high
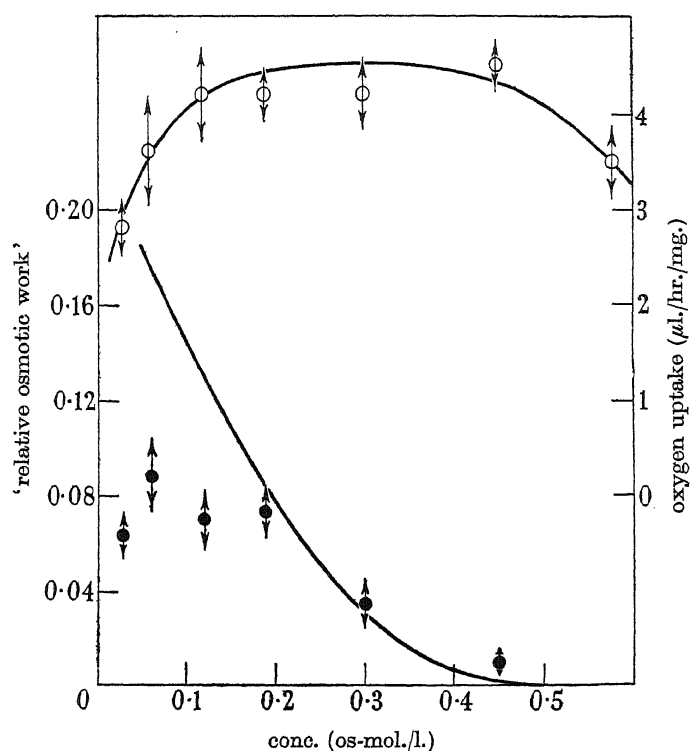
FIGURE 13. 'Relative osmotic work' required to maintain observed water content of slices in solutions of different concentrations. ● osmotic work. ◑ oxygen uptake. Arrows show ± S.D. Numbers of experiments in table 1. ——— Osmotic work required to prevent swelling.

temperature coefficient, and values of Q_{10} from 2 to 4 are not uncommon (Brooks & Brooks 1941; Davson & Danielli 1943). Hence the amount of osmotic work needed to keep the cells to a given size might diminish more rapidly with falling temperature than the energy available from respiration. In order to derive the theoretical curve in figure 7, the respiratory rates at temperatures from 16 to 38° C were read off the smooth curve drawn through the experimental points in figure 6. The amounts of osmotic work corresponding to these respiratory rates were then found from the line drawn through the points in figure 12, and divided by the temperature corrections required in equation (10) on the assumption that the temperature coefficient of the permeability constant k was 3·2. The results were assumed to be the amounts of available osmotic work which would correspond at 38·5 °C to the actual respiratory

rates at the temperatures in question, and the percentages of water which could have been maintained in the slices by these rates of osmotic work were read off the curve for $m_o = 0.30$ in figure 11 and plotted against the temperatures in figure 7. It will be seen that the theoretical curve predicts both the absolute values which were found in the experiments, and also the tendency for the amount of water in the slices to pass through a minimum at about 30° C.

(4) *The absolute value of osmotic work calculated from the theory*

The theoretical treatment which has been described could not be regarded as plausible if it led to ridiculous results for the absolute value of the osmotic work performed in the slices. The order of magnitude in cal./mg. of tissue per hour was therefore calculated approximately by two independent methods, for the water content of 77·5 % which was maintained in isotonic solutions by an oxygen uptake of 4·2 μl./hr./mg. of tissue. Assuming an energy yield of 5 kcal/l. of oxygen, the respiration could have provided $4.2 \times 10^{-6} \times 5$ kcal., $= 0.02$ cal./hr./mg. of tissue.

Method (i). By substituting assumed values of k and A in equation (10): Lucké, Hartline & Ricca (1939) quoted values of the order of 1 μ^3/μ^2/min./atm. difference of osmotic pressure for the permeability to water of mammalian erythrocytes, leucocytes and fibroblasts. These were measured at 24 to 28° C, so that if Q_{10} was of the order of 3, $3\mu^3/\mu^2$/min./atm. should be a reasonable figure for the permeability of cells at 38·5° C. When it was converted into suitable units to use in equation (10), this gave $k = \frac{3}{18} \times 22.4 \times 60 \times 10^{-12} = 2.2 \times 10^{-10}$ mols of water diffusing across each μ^2 of surface per hour under a concentration difference of 1 os-mol/l. A spherical cell 20μ in diameter would have a volume of about $4000\mu^3$, and a surface area of about $1200\mu^2$. For a cubical cell with a side 15μ in length the volume would be $3375\mu^3$ and the area $6 \times 225 = 1350\mu^2$. If now 1 mg. of tissue contained 26 % of extracellular fluid, and therefore 0·74 mg. of cells of average volume $3500\mu^3$ and surface area $1300\mu^2$, the total number of cells in the 1 mg. of tissue would be $\frac{0.74 \times 10^9}{3500} = 2 \times 10^5$ approx., and their total surface would be $2 \times 10^5 \times 1300 = 2.6 \times 10^8\mu^2$. Further, since the value of S_c corresponding to water content of 77·5 % was 1·10 (figure 9), $m_i = 0.52/1.10 = 0.47$. The substitution of these figures in equation (10) gave W_s (in cal./hr./mg. of tissue)

$$= \frac{kART}{3025}(m_i - m_o)^2$$

$$= \frac{2.2 \times 10^{-10} \times 2.6 \times 10^8 \times 1.99 \times 311}{3025}(0.47 - 0.30)^2$$

$$= 0.0003 \text{ cal./hr./mg.}$$

This is only about 1·5 % of the energy available from respiration, but the estimate depends so much upon the values of A and k that it might be in error by a factor of 10 or even more.

Method (ii). Another estimate was made more simply by using equation (9). In mammalian kidneys it is usual to find glomerular filtration rates of the order of

half the volume of kidney tissue per minute, whilst the urine volumes are only a small fraction of this. It follows that water can traverse the renal cells at rates of the order of 30 times the cell volume per hour. Assuming a similar rate of exchange in the steady state, the cells in 1 mg. of tissue which maintained a water content of 77·5 % in a 0·30 osM solution would have had to expel 30 mg. or $30/18 \times 10^{-3} = 1·7 \times 10^{-3}$ mols of water per hour. If the internal concentration was 0·47 osM as above, the amount of energy required, from equation (9), would be

$$\frac{nRT}{55}(m_i - m_o) = \frac{1·7 \times 10^{-3} \times 1·99 \times 311}{55}(0·47 - 0·30)$$

$$= 0·0032 \text{ cal./mg. of tissue/hr.,}$$

or about 15 % of the energy available from respiration.

Probably the only conclusion that can be drawn from these extremely rough estimates is that the mechanism which has been outlined could account for the observations without requiring an impossible amount of energy. Moreover, if the ionic differences between intracellular and extracellular fluids were maintained as steady states by a process requiring an expenditure of energy of the same order, then unless the efficiency of the transporting mechanisms were greater than about 30 % (Franck & Mayer 1947), a substantial part of the resting oxygen consumption of the tissues would be required for maintaining differences in composition between the cells and their surroundings. A further interesting point which arises out of the second method of calculation above is that if the cells possessed a functional polarity, so that water diffused in at one pole and was extruded at the other, the normal energy of maintenance would be sufficient to account for the rates of tubular reabsorption found in the intact kidney. In fact, the cells of the proximal tubules could perform their specific function of reabsorbing water, and also maintain their own normal water content by the same operation. The secretory activity of cells of this kind might then be regarded as differing quantitatively rather than qualitatively from the activity of most resting cells. Finally, the theory of the water balance of the body as a steady state could account for all the observations which could be explained on the classical theory of osmotic equilibrium as well as for the results of the present experiments, and it is more in accordance with current ideas of the 'Dynamic state of body constituents'.

REFERENCES

Adolph, E. F. 1947 *Ann. Rev. Physiol.* 9, 381.

Bacon, J. S. D. & Bell, D. J. 1948 *Biochem. J.* 42, 397.

Bertalanffy, A. von 1950 *Science*, 111, 23.

Blumenthal, R. 1927 *Biol. Bull. Woods Hole*, 52, 313.

Borsook, H. & Winegarden, H. M. 1931 *Proc. Nat. Acad. Sci., Wash.*, 17, 3.

Brooks, S. C. & Brooks, M. M. 1941 *The permeability of living cells. Protoplasma Monographien*, 19. Berlin: Borntraeger.

Cohen, P. P. In Umbreit, W. W., Burris, R. H. & Stauffer, J. F. 1945 *Manometric techniques*, p. 75. Minneapolis: Burgess.

Conway, E. J. & Fitzgerald, O. 1942 *J. Physiol.* **101**, 86.

Conway, E. J., Fitzgerald, O. & MacDougald, T. C. 1946 *J. Gen. Physiol.* **29**, 305.

Davson, H. & Danielli, J. F. 1943 *The permeability of natural membranes.* Cambridge University Press.

Franck, J. & Mayer, J. E. 1947 *Arch. Biochem.* **14**, 297.

Gaudino, M. & Levitt, M. E. 1949 *J. Clin. Invest.* **28**, 1487.

Glasstone, S. 1947 *Thermodynamics for chemists.* New York: van Nostrand.

Gömöri, P. & Molnár, S. 1932 *Arch exp. Path. Pharmak.* **167**, 459.

Gömöri, P. & Frenreisz, St. 1937 *Acta med. Scand.* **92**, 497.

Hill, A. V. 1930a *Trans. Faraday Soc.* **26**, 667.

Hill, A. V. 1930b *Proc. Roy. Soc. A*, **127**, 9.

Hill, A. V. 1930c *Proc. Roy. Soc. B*, **106**, 477.

Hill, A. V. 1931 *Adventures in biophysics.* Oxford University Press.

Kitching, J. A. 1938 *Biol. Rev.* **13**, 403.

Kitching, J. A. 1948 *J. Exp. Biol.* **25**, 406.

Krebs, H. A. 1950 *Biochim. Biophys. Acta,* **4**, 249.

Lucké, B., Hartline, H. K. & Ricca, R. A. 1939 *J. Cell. Comp. Physiol.* **14**, 237.

Opie, E. L. 1949 *J. Exp. Med.* **89**, 185.

Peters, J. P. 1935 *Body water.* Springfield: Thomas.

Peters, J. P. 1937–38 *Harvey Lectures,* **33**, 112.

Peters, J. P. 1944 *Physiol. Rev.* **24**, 491.

Robinson, J. R. 1939 *Proc. Roy. Soc. A*, **170**, 519.

Robinson, J. R. 1949 *Biochem. J.* **45**, 68.

Robinson, J. R. 1950. *Nature*, **165**, 730.

Sabbatani, L. 1901 *J. Physiol. Path. gén.* **3**, 939.

Siebeck, R. 1912 *Pflüg. Arch. ges. Physiol.* **148**, 443.

Stern, J. R., Eggleston, L. V., Hems, R. & Krebs, H. A. 1949 *Biochem. J.* **44**, 410.

Action of nitrogen trichloride on certain proteins
I. Isolation and identification of the toxic factor

By H. R. Bentley, E. E. McDermott, T. Moran, J. Pace and J. K. Whitehead

Research Association of British Flour-Millers, Cereals Research Station, St Albans

[Plate 20]

When proteins in which the content of methionine is relatively high are treated with nitrogen trichloride they become toxic to certain animals. This effect is observed, for example, with zein, gluten and egg albumen. If these treated proteins are digested, either enzymatically or by partial acid hydrolysis, the toxic property persists in the hydrolysates.

Methods are described by which such hydrolysates have been fractionated, and the toxic component isolated as a pure crystalline substance. The toxic factor derived from each of the three proteins, zein, gluten and egg albumen, is found to be the same chemical substance.

Data are given which establish that it is a derivative of the amino-acid methionine with the molecular formula $C_5H_{12}N_2O_3S$, and the structural formula

$$\text{HOOCCHCH}_2\text{CH}_2 \underset{\underset{\text{CH}_3}{\diagup}\diagdown_{\text{NH}}^{-}}{\overset{\overset{\text{O}}{\diagup}}{\underset{\text{S}}{\diagdown}}}$$
$$\underset{\text{NH}_2}{|}$$
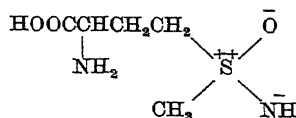
This molecule exhibits diastereo-isomerism, and it is probable that the various isomers have differing biological activity.

INTRODUCTION

Mellanby (1946) showed that when dogs are fed a diet rich in flour which has been treated with the flour-improver known commercially as 'agene' (in which the essential active reagent is nitrogen trichloride) they develop the condition known as canine hysteria. The more obvious symptoms of this disorder are bouts of hysterical barking and aimless running, followed in severe cases by epileptiform fits. It was soon established that similar symptoms develop in ferrets (Mellanby 1947), cats (Newell, Erickson, Gilson, Gershoff & Elvehjem 1947), and rabbits (Radomski, Woodward & Lehman 1948) when they are fed with flour which has been treated with nitrogen trichloride.

Mellanby's observations raised the problem of the nature of the reaction of nitrogen trichloride with flour, and of how the toxic property is produced. As a first approach to these questions it was desirable to ascertain whether the production of toxic material by the use of nitrogen trichloride is specific to wheaten flour, whether particular components of flour are specifically implicated or whether the reaction is of a more general character and not confined to wheaten flour. Experiments reported by Moran (1947) provided conclusive data on these points and thereby facilitated the development of a rational approach to the problem of the isolation and nature of the toxic factor. In this publication it was shown that the toxic factor is produced by the interaction of nitrogen trichloride with the protein fraction of wheaten flour, gluten, and further that the proteins zein and casein also become toxic when treated with the gas. These observations were extended in a later paper (Bentley, Booth, Greer, Heathcote, Hutchinson & Moran 1948), when it was shown that the proteins egg albumen, haemoglobin and rice protein become toxic on treatment with nitrogen trichloride, but that arachin and kerateine, although interacting readily with nitrogen trichloride, remain non-toxic. The essential reaction with nitrogen trichloride is thus of a fairly general character, since toxic material is obtained from a range of proteins of both animal and vegetable origin. On the other hand, it is not common to all proteins. From a consideration of the amino-acid composition of the various proteins, Bentley *et al.* (1948) suggested a possible explanation of these results. The proteins which readily become toxic are relatively rich in methionine, while those which remain non-toxic contain little or none of this amino-acid. Other amino-acids which may be expected to react with nitrogen trichloride are not so significantly distributed between the two groups of proteins. It was therefore suggested that within the protein the unit which reacts with nitrogen trichloride to form the toxic factor contains a methionine residue as an essential part of its structure. Further support for this suggestion was found in the fact that when the methionine residues in zein and casein are modified by oxidation of the proteins with hydrogen peroxide, in appropriate solvents, the oxidized proteins after interaction with nitrogen trichloride are not toxic. On the other hand, it was observed that neither methionine *per se* nor an acid hydrolysate of casein or zein became toxic when treated with nitrogen trichloride, though both readily absorb it.

These investigations suggested strongly that the problem of determining the nature of the toxic factor was essentially one of isolating a modified peptide unit or

a modified amino-acid from a suitable protein treated with nitrogen trichloride. They also provided some ground for presuming that methionine is implicated in the toxic grouping. With these suggestions as a guide, experiments were initiated with the aim of isolating the toxic factor from zein treated with nitrogen trichloride. This work has led to the isolation of the toxic factor as a pure crystalline substance and to its characterization as a derivative of methionine. These results have already been briefly reported in preliminary notes (Bentley, McDermott, Pace, Whitehead & Moran 1949a, b, 1950a), and the purpose of the present communication is to give a fuller description of the work carried out. Following its isolation from zein, the same toxic crystalline substance has been isolated from gluten and crystalline egg albumen treated with nitrogen trichloride, and this work is described later in this paper.

The isolation of the toxic factor from zein and its characterization as a methionine derivative has recently been confirmed independently by Dr L. Reiner and his colleagues in the U.S.A. (Reiner, Misani & Weiss 1950). A paper by Campbell, Work & Mellanby (1950) also reports the isolation, from nitrogen trichloride-treated flour, of a crystalline toxic factor which appears to be identical with the methionine derivative reported in our earlier notes.

ISOLATION OF THE TOXIC FACTOR, AS A PURE CRYSTALLINE SUBSTANCE, FROM ZEIN TREATED WITH NITROGEN TRICHLORIDE

Assay of toxity

Throughout the investigation fractions and samples of material have been tested for toxicity by feeding them to rabbits and, occasionally, checks have been made by feeding to dogs. Individual rabbits vary considerably in their sensitivity to the toxic factor, and a large number of animals would be required to establish quantitatively the significant minimum toxic dose at the various stages of the isolation process. A simple technique has, however, proved adequate as a guide to the efficacy of processes designed to concentrate the toxic factor, and has provided a semi-quantitative assessment of the toxicity of different fractions. This has consisted of feeding any given sample to each of a group of three rabbits and taking a result to be positive if at least two of the three develop typical symptoms within 48 hr. On this basis, as a rough estimation, a toxic dose is defined as the weight of material required to produce typical running fits and/or convulsions within 48 hr. in rabbits of approximately 1 kg. weight. All soluble fractions have been administered to rabbits, in the minimum quantity of water, by means of a stomach tube.

In the course of the work it was observed by Heathcote (1949) that fractions toxic to rabbits are also toxic to the micro-organism *Leuconostoc mesenteroides* P. 60, and in view of the high sensitivity of this organism to the toxic factor, its use is of great value as an indicator of toxicity on occasions when only small quantities of material are available.

Preparation of nitrogen trichloride

A convenient laboratory method of preparing nitrogen trichloride is to aerate an aqueous solution of chlorine in ammonium chloride or ammonium sulphate solution.

Equal volumes of freshly prepared saturated chlorine water and of 2 % ammonium sulphate are mixed and kept below 10° C in an amber-coloured bottle in the dark for about 30 min.

With solutions at pH 4·4 or less the ammonium derivative is almost entirely nitrogen trichloride with very little mono- or dichloramine. On aerating the solution the effluent gas stream contains air, water vapour, nitrogen trichloride vapour and some chlorine. It is this gas mixture which is used commercially in 'agenizing' flour, and it has been used in the present investigation for the treatment of proteins.

Preliminary experiments established that there is no production of the toxic factor when flour is treated with chlorine.

Treatment of zein with nitrogen trichloride

Coarse zein powder (obtained from Corn Products Co. Ltd) is ground between steel balls in a ball-mill to a fineness sufficient to pass a no. 14 silk. Batches of 2 kg. of the fine powder are agitated in a wooden box rotating about a hollow spindle through which nitrogen trichloride vapour is gently blown by a current of air from the container in which it is prepared. Agitation of the powdered zein in an atmosphere of nitrogen trichloride vapour is continued for several hours until the protein has absorbed approx. 5·5 mg. NCl_3 per gram. The resulting zein is now toxic— a toxic dose for a rabbit being about 5 g. and for a dog 25 g.

Enzymatic digestion of the toxic zein

Preliminary experiments established that when proteins, rendered toxic by nitrogen trichloride, are digested by proteolytic enzymes the toxic property is retained in the digestion mixture. Digestion thus provides a useful method of obtaining toxic material from zein in a water-soluble form, and combined with dialysis to remove undigested protein and peptides of high molecular weight, it is a convenient first step in concentrating the toxic factor. One gram of pancreatin powder (Armour and Co. Ltd triple strength) is added to a suspension of 100 g. of treated protein in 250 ml. of water containing 5 ml. of ammonia (density 0·88). The mixture is contained in a stoppered 1 l. conical flask, and is preserved by the addition of a little toluene. The mixture is digested at 37° C for 3 days with frequent shaking, a further 1 g. of pancreatin being added on the second day. At the end of this digestion period the contents of the flask are transferred to Cellophane tubing and dialyzed against distilled water (protected with toluene) for a total of 7 days, at room temperature, the dialysate being replaced by fresh water on the first and third days. From a batch of 2 kg. of zein the volume of the combined dialysate is about 10 l. and contains about 450 g. of solids. This volume is reduced by distillation under reduced pressure on a water-bath until the final solution contains 10 to 12 % solids (w/v).

Alternatively, soluble toxic material may be obtained directly from the enzymatic digest by filtration, and the dialysis step omitted. While this procedure affords a considerable saving of labour it has not, in our experience, proved consistently satisfactory and the dialysis method is preferred.

The aqueous dialysate is toxic at the level of 1·5 g. solids per rabbit.

Solvent extraction

The aqueous dialysate, of 10 to 12 % solids content, is extracted at 80° C with *n*-butanol water saturated at this temperature; extraction is first carried out with a volume of butanol equal to that of the aqueous solution, and this is followed by two further extractions with one-half the same volume. These extractions remove, in the butanol phase, approximately one-third of the total solids without significant loss of toxicity from the aqueous phase, although toxic material is extracted to an increasing extent if the number of butanol extractions is extended beyond three.

The aqueous phase is freed from butanol by concentration under reduced pressure on the steam bath to two-thirds its original volume and is then extracted at room temperature by shaking it, in separating funnels, with crystalline phenol (40 g. phenol per 100 ml. solution). The lower (phenol-rich) phase is separated, diluted with an equal volume of water and the phenol removed by repeated extraction with ether. The aqueous phase remaining after removal of the phenol is concentrated under reduced pressure. This material is toxic at the level of 0·6 g. solids per rabbit.

Repeated attempts have been made to fractionate directly the constituents of the solution, derived from the phenol phase, by means of partition chromatography. A preliminary survey indicated, using paper chromatograms, that *n*-butanol, phenol and collidine each effect a separation of the material into discrete bands, and operations were therefore continued on a larger scale. Using water-saturated *n*-butanol with columns of silica gel (Isherwood 1946) or starch (Stein & Moore 1948), toxicity is found to be mainly concentrated in a stationary or very slowly moving band, while faster-moving material, which may be easily eluted, is non-toxic; with phenol on columns of silica gel toxic material appears in the fastest-moving band.

Invariably, however, it has been found that fractionation of material of this degree of purification on partition chromatograms is unsatisfactory, as there is a considerable loss of toxic material. In this material there are present free amino-acids and a range of peptides of varying molecular weight. It is possible that the toxic factor grouping is distributed among these different molecular species, and that therefore there is no concentration of all the toxic factor in one discrete band on the partition chromatograms. Equally it may be that a proportion of the toxic material is irreversibly adsorbed on silica gel. After many attempts to improve the recovery of toxic material without success it was decided to abandon partition chromatography with material of this degree of purification. It was considered that a more profitable approach could be made if the material from the phenol phase were further reduced, by acid hydrolysis, to amino-acids and small peptides. A study of the stability of the material to various conditions of acid hydrolysis was therefore made, and it was found that the toxic factor is remarkably stable to boiling acid. On boiling with hydrochloric acid (initially 8 N) under reflux for 3 hr. there is no apparent loss of toxicity, but with more prolonged boiling there is a gradual destruction of the toxic factor. These observations led to the adoption of acid hydrolysis as a step preliminary to further purification.

Acid hydrolysis and further solvent extraction

A solution of the phenol-soluble solids (10 % w/v) in hydrochloric acid (initially 8 N) is boiled under reflux for 3 hr., and the bulk of excess free acid is then removed by repeated evaporation to dryness under reduced pressure on the steam bath. The dry residue of mixed hydrochlorides is then redissolved in sufficient water to give a solution containing 10 % (w/v) solids which is heated to 80° C on the water-bath, and then extracted five times with half its own volume of *n*-butanol heated to 80° C and water-saturated at this temperature. The aqueous phase is again concentrated under reduced pressure to a content of 10 % (w/v) solids, and then mixed with a quantity of activated charcoal (Sutcliffe and Speakman no. 110, previously boiled for 2 hr. with 20 % acetic acid, washed with a large volume of water and dried) equivalent to 25 % of the solids in solution. After half an hour the solution (now clear and colourless) is filtered from the charcoal, when it contains material toxic at the level 0·3 to 0·4 g. solids per rabbit (as free base).

Acetone precipitation

The clear aqueous solution (100 ml.), from the charcoal treatment, is mixed with dry acetone (1710 ml.) in a stoppered 2 l. conical flask and kept overnight in the refrigerator; the supernatant liquid is then decanted and any remaining excess acetone allowed to evaporate at room temperature from the oily precipitate. The solids in the precipitate normally constitute approximately 15 % of the charcoal-treated fraction and are toxic at the level of approximately 150 mg., which represents a threefold concentration of toxicity. If the acetone-insoluble fraction represents a considerably higher proportion (20 to 25 %) of original solids it is redissolved in water and the precipitation repeated with a further quantity of dry acetone. While acetone precipitation appears to be the least efficient step in the series of fractionations (since considerable amounts of the toxic factor are retained in the soluble fraction), it does, however, seem to be essential for the further purification of the toxic factor by the methods we have employed. Attempts to utilize the acetone-soluble fraction directly for further purification have met with little success. It is not clear why this should be so, but it may be significant that the soluble fraction contains a very much higher proportion of the leucines than the insoluble fraction. These amino-acids overlap the band containing toxic factor on the ion-exchange resin column used at a later stage.

Electrodialysis

A solution of the material insoluble in 95 % acetone (at a concentration of approximately 1 % w/v) in water is next electrodialyzed in a three-compartment cell (Albanese 1940), using carbon plate electrodes. The electrode compartments each have a capacity of approximately 200 ml. and are cooled by circulation of water through a grid; the centre compartment has a similar capacity, and its contents are stirred by a stream of air bubbles. The cathode membrane is vegetable parchment (British Vegetable Parchment Mills, Grade no. 40.44) and the anode membrane formolized gelatin supported on linen. A group of four cells is wired in parallel to

a source of 110 V d.c. Each cell carries in series a milliammeter and a 100 W lamp. Initially the centre compartment is filled with solution adjusted to pH 7·0 with ammonia (density 0·88), and the electrode compartments are filled with distilled water. Electrodialysis is continued until the current falls to 100 mA, when the anolyte and catholyte are withdrawn and replaced by water, the pH of the centre compartment is again adjusted to 7·0 and electrodialysis resumed until the current falls below 25 mA. The combined catholytes, after concentration *in vacuo*, are now placed in the centre compartment and re-electrodialyzed, the pH being kept at about 6·0 by the addition of ammonia. The pH of the contents of the centre compartment eventually steadies at about 5·8 to 6·0, and the process is stopped when the current falls below 10 mA. It is important to ensure in this operation the maximum transference of histidine to the catholyte. The combined centre compartment fractions of pH 5·8 to 6·0 are retained for further purification and the final catholyte is rejected.

Ion-exchange fractionation

The toxic fraction from electrodialysis is now fractionated further on a column of Zeo-Carb 215 (Permutit Co. Ltd). The coarse resin is ground in a hammer mill, and the fraction passing a no. 10 silk but retained by a no. 14 silk is employed. The resin (20 g.) is washed thoroughly by decantation with water and poured as a thin slurry into a length of glass tubing (8 mm. internal diameter, constricted to a narrow orifice at the lower end) and allowed to settle to form a column. After conditioning by percolation alternatively twice or three times with hydrochloric acid (2 N) and ammonia (N), the column contracts to a final length of approximately 110 cm. The column is used in the acid form, after free acid has been removed by washing with water.

The mixed amino-acids (3·0 to 4·0 g.) are applied as a 5 % solution in water to the column and washed in thoroughly with water (3 × 10 ml.). The absorbed amino-acids are clearly visible as a paler brown band on freshly prepared resin and are now displaced with dilute ammonia (0·20 N), the ammonia boundary appearing as the leading edge of a dark brown band. As soon as amino-acids appear in the eluate the latter is collected in small fractions (approximately 5 ml.) until it becomes free from ninhydrin-reacting substances shortly after the point of ammonia 'breakthrough'. Samples of these fractions are spotted in series on sheets of filter-paper (Whatman no. 1) and developed overnight by downward irrigation with water-saturated phenol in conventional apparatus (Consden, Gordon & Martin 1944; Dent 1948). A photograph of a typical series is reproduced in figure 1, plate 20. Shortly after the appearance of proline (recognized as the typical yellow spot, R_F 0·8) in the eluate, fractions are found to contain a new spot (R_F 0·7) which persists to the end of the run; this spot may be recognized by its typical sequence of colour changes from yellow through brown to purple when the paper chromatogram is heated gently after spraying with ninhydrin (0·1 % w/v in n-butanol), although in the end-fractions this effect may be masked by the stronger grey-purple histidine coloration which is superimposed on it. All fractions containing the new spot, R_F 0·7, are combined for further purification. It is possible, with experience, to
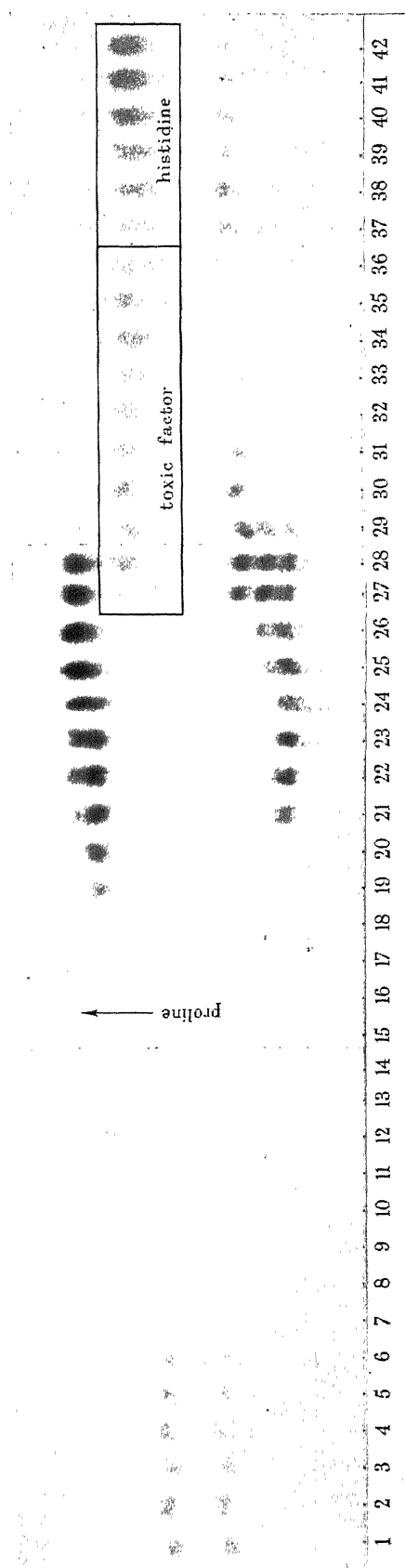
FIGURE 1. Photograph of filter-paper chromatograms showing the sequence in which ninhydrin-reacting substances, from proline onwards, emerge from the ion-exchange column 'Zeo-Carb 215', and indicating the position of the fractions containing the toxic factor. The paper chromatograms are obtained using Whatman no. 1 filter-paper irrigated with phenol.

reduce the labour of spotting every ninhydrin-reacting fraction by rejecting that portion of the amino-acid band which is eluted before proline appears in the eluate; with a column giving a satisfactory separation, collection in small fractions can be delayed until the eluate reacts pure yellow with ninhydrin on a filter-paper.

The 'post-proline' band contains material toxic at the level of 8 to 10 mg.

Partition chromatography

In addition to the ninhydrin-reacting substances with R_F 0·7 in phenol the toxic post-proline fraction from the ion-exchange column contains several other ninhydrin-reacting substances, namely, proline, the methionine-leucines group and a group of unidentified substances with R_F values in the range 0·2 to 0·4. This fraction is now separated further by means of partition chromatography.

Partition columns are prepared in glass tubes (2·5 cm. diameter, 50 cm. length) constricted to approximately 5 mm. at the lower end; a perforated porcelain plate covered with a thin layer of glass-wool at the base of the tube serves to support the column. Kieselguhr (British Drug Houses Ltd acid-washed, 75 g.) mixed with water (50 ml.) is stirred with water-saturated phenol (500 ml.), the slurry poured into the glass tube and the column packed by gentle pressure with a perforated aluminium plate according to the method demonstrated by Randall & Martin (1949).

The completed column is approximately 33 cm. long, and its preparation occupies 2 days; excess unpacked kieselguhr is withdrawn by means of a pipette from the surface of the column and the inner wall of the tube above the column washed down with a little water-saturated phenol. Excess phenol is allowed to drain into the column, and this is followed by a solution of dry toxic fraction (150 mg.) in liquefied phenol B.P. (3 ml.) which is then washed in with water-saturated phenol. The column is now developed with water-saturated phenol, and ninhydrin-reacting material appears when approximately 100 ml. of eluate has been collected. The eluate is now collected in small fractions (5 ml.) and the separation followed by means of paper chromatography exactly as for the ion-exchange column. Fractions containing material with R_F 0·7 are combined, diluted with water and freed from phenol by extraction with ether.

The toxic fraction (toxic dose 3 to 4 mg.) from the partition column gives a single ninhydrin-reacting spot on paper chromatograms in phenol. In collidine and in aqueous acetone it is resolved into three distinct spots; one of these is very faint, another is identified as being due to histidine. The mixture gives an intense Pauly reaction.

Ion-exchange fractionation

A solution of toxic material from the previous stage (approximately 1 % w/v in water) is applied to a column of Zeo-Carb 215 (diameter 0·55 cm., length 20 cm.), prepared and conditioned as described previously. The material is washed into the column with water (5 ml.) and then eluted with dilute ammonia (0·15 N). When ninhydrin-reacting material appears in the eluate this is collected in small fractions (1 ml. or less). The fractions first eluted are at pH 5·8 and give a negative Pauly reaction, the pH subsequently rises abruptly through pH 6·5 to 7·0 when histidine

appears in the eluate, and these fractions give the Pauly reaction which remains positive to the end. The fractions giving no Pauly reaction are combined; on paper chromatograms they give a single spot in collidine and acetone. The combined

TABLE 1. SUMMARY OF THE STAGES IN THE ISOLATION OF THE TOXIC FACTOR

sequence of operations	weight of solids (g.) in toxic fractions	toxic dose (g.) for rabbits	% yield on original toxicity	% yield on original weight
zein treated with 5·5 *mg.* $NCl_3/g.$ *protein*	2000	5	—	—
pancreatin digestion and dialysis				
dialysate	450	1·5	75	22·5
butanol extractions				
aqueous phase	300	1·0	—	—
phenol extraction				
phenol phase	150	0·6	62	7·5
acid hydrolysis increases weight to 200 g.	200	0·8	—	—
butanol extraction				
aqueous phase	100	0·4	62	5·0
acetone precipitation				
insoluble	12	0·15	20	0·6
electrodialysis				
combined material from centre compartment	3·8	0·05	19	0·2
ion-exchange column 1				
end fraction	0·6	0·008*	19	0·03
kieselguhr partition column				
material R_F (*phenol*) 0·7	0·20	0·003	17	0·010
ion-exchange column 2				
material giving negative Pauly reaction	0·12	0·002	15	0·006
aqueous ethanol crystallization				
crystalline factor	0·080	0·002	10	0·004

* In our earlier note the toxic dose at this stage is given as 0·004 g. This is due to the fact that then only material from the centre compartment of the first electrodialysis was used. There was no re-electrodialysis of the catholyte.

fractions are evaporated to dryness under reduced pressure on the steam bath, the dry residue taken up in the minimum quantity of water and absolute ethanol added up to a concentration of approximately 60 to 80%. Crystallization commences almost at once and is complete after several hours in the refrigerator. The crystalline material is collected, washed with ethanol and dried *in vacuo*. It is toxic at the level of 2 mg.

Control

As a control experiment 2 kg. of 'unagenized' zein have been processed by methods exactly similar to those described above. The fractions obtained at each stage prior to the second ion-exchange fractionation are similar to those found with 'agenized' zein. Examination by paper chromatography of the control fractions after the first ion-exchange fractionation and after partition chromatography show in each case a fraction with R_F (phenol) 0·7 as before, and in all other respects the paper chromatograms are apparently identical. It is found, however, that the spot with R_F 0·7 does not exhibit the characteristic series of colour changes when heated on paper with ninhydrin; only the grey-purple coloration due to histidine is observed. The fraction R_F 0·7 (phenol) from the kieselguhr column contains very little material, and on the second ion-exchange column the fraction first eluted is at pH 6·5 and gives a strongly positive Pauly reaction due to histidine. There is no material present corresponding to the toxic crystalline material obtained from 'agenized' zein.

The fractionation of total solids and of approximate toxicity at each stage of the purification of the crystalline substance derived from 'agenized' zein is conveniently summarized in the flow-sheet given in table 1.

ISOLATION OF THE CRYSTALLINE TOXIC FACTOR FROM GLUTEN TREATED WITH NITROGEN TRICHLORIDE

Gluten is prepared by washing out from 'unagenized' flour. It is then partially dried under acetone, pulverized under acetone in a Waring Blendor and finally washed well with acetone and ether and air-dried. The resulting powder is sieved through a no. 8 silk (aperture 0·19 mm.) and is then treated with nitrogen trichloride as described above for zein.

1200 g. of gluten treated with nitrogen trichloride are digested with pancreatin and the digestion mixture dialyzed for 4 days. The solid content of the dialysate is about 370 g. and contains approximately 180 toxic doses of the factor.

The dialysate is then fractionated by exactly the same procedure as described for zein and yields finally 10 mg. of crystalline toxic factor.

The factor is toxic to rabbits at the 2 mg. level, is toxic to *L. mesenteroides* and is identical with that obtained from zein when examined by paper chromatography using phenol, acetone and collidine as developing solvents.

ISOLATION OF THE CRYSTALLINE TOXIC FACTOR FROM EGG ALBUMEN TREATED WITH NITROGEN TRICHLORIDE

An aqueous solution of crystalline egg albumen is prepared from the white of eggs by the method of Hopkins (see Cole 1923). From the solution dry, undenatured

crystalline egg albumen is obtained by carefully vacuum 'freeze-drying'. (This operation was carried out by Dr Gane at the Low Temperature Research Station, Cambridge, to whom we gratefully acknowledge our thanks.)

Using the freeze-dried albumen, a preliminary investigation was made of the absorption of nitrogen trichloride by the native crystalline protein compared with the absorption by the protein when it is denatured by heat. The native albumen when fully saturated absorbs approximately 5·0 mg. of nitrogen trichloride per g. of protein, and the product is about as toxic as zein which has been treated with nitrogen trichloride at the same level. Egg albumen denatured by heating at the isoelectric point absorbs at full saturation about 110 mg. nitrogen trichloride per g. of protein, and the product is toxic to a dog at the 4 g. level. The increase in absorption of nitrogen trichloride when the protein is denatured is presumably due largely to the 'unmasking' of SH groups, but, in addition, there appears to be a definite increase in the availability of methionine residues for interaction with nitrogen trichloride in the denatured as compared with the native protein.

Denatured egg albumen is used for the isolation of the toxic factor. A solution of native albumen is adjusted to about pH 4·5 with a few drops of glacial acetic acid and then slowly heated with constant stirring to about 80° C. The flocculated denatured protein is collected on a Buchner funnel, washed with water and then thoroughly with acetone. It is then air-dried. The procedure for the isolation of the toxic factor differs in some particulars from that used with zein and gluten.

8 g. of denatured egg albumen are treated with nitrogen trichloride until there is no further absorption. Pancreatic digestion and dialysis are omitted, and the toxic protein is directly hydrolyzed by boiling under reflux with hydrochloric acid (initially 8 N) for 8 hr. The hydrolysate is extracted five times with hot water saturated n-butanol, and the aqueous phase from these extractions is treated with charcoal. Most of the free hydrochloric acid is then removed by repeated evaporation under reduced pressure. Further chloride is removed by adjusting the solution to about pH 3·5 by the cautious addition of freshly prepared silver oxide. Precipitated silver chloride is removed by filtration and the filtrate is then electrodialyzed. During electrodialysis the pH of the centre compartment is carefully maintained at about pH 6·0. The material from the centre compartment is then fractionated on ion-exchange and kieselguhr columns as described above for zein.

The toxic factor obtained from egg albumen has been identified with that from zein by paper chromatography, using phenol, acetone and collidine as solvents, and by toxicity assays with L. mesenteroides.

CHARACTERIZATION OF THE CRYSTALLINE TOXIC FACTOR

Toxicity

The recrystallized material is toxic at the level of approximately 2 mg. orally in rabbits. Shortage of material has precluded extensive feeding trials with dogs, but in a single experiment a 10 kg. dog given an aqueous solution of 20 mg. of crystalline material by stomach tube developed all the symptoms of hysteria, including

running and barking fits, in less than 18 hr., followed by violent convulsive fits. The minimal dose to produce hysteria is probably much less than 20 mg.

Chemical properties

The toxic factor crystallizes from aqueous ethanol as small colourless needles which darken at 225 to 230° C and melt with decomposition at 234° C. It is freely soluble in water, sparingly soluble in methanol and ethanol and not appreciably soluble in non-polar solvents. Under conditions described in detail elsewhere (Bentley & Whitehead 1950), the toxic factor has R_F values 0·71 in water-saturated phenol and 0·43 in aqueous acetone (40 % water v/v) on Whatman no. 1 paper. In table 2 these values and the corresponding values for methionine, methionine sulphoxide and methionine sulphone are compared.

TABLE 2. R_F VALUES, USING WHATMAN NO. 1 FILTER-PAPER, WITH PHENOL AND AQUEOUS ACETONE (40 % WATER V/V). THE IRRIGATION OF THE PAPERS WITH PHENOL IS DOWNWARDS AND WITH AQUEOUS ACETONE UPWARDS

	phenol	acetone
methionine	0·80	0·53
methionine sulphoxide	0·78	0·42
methionine sulphone	0·69	0·57
toxic factor	0·71	0·43

Analysis

Analysis of different batches of crystals correspond most closely with an empirical formula $C_5H_{12}N_2O_3S$. (Found: C 33·8, 34·2; H 6·7, 6·4; N 15·6, 15·4; S 17·3, 17·3. $C_5H_{12}N_2O_3S$ requires C, 33·3; H, 6·7; N, 15·6; S, 17·8 %.) Carbon and hydrogen determinations are by combustion (Drs Weiler and Strauss, Oxford). Nitrogen determinations are by the micro-Kjeldahl method (HgO/K_2SO_4 catalyst). Sulphur determinations are by combustion in a stream of hydrogen (Olley 1945) and were carried out by Miss M. Corner of the Chemical Research Laboratory, Department of Scientific and Industrial Research, and by Mr W. J. S. Pringle in our own laboratories. The micro-Carius sulphur determination gives consistently low results (14·9, 9·5, 9·6 %), an effect which has already been noted in the case of methionine and related compounds (Callan & Toennies 1941). Molecular weight determination by Barger's method in water gives a value of approximately 200; $C_5H_{12}N_2O_3S$ requires 180.

The presence in the molecule of one α-amino-carboxylic acid grouping is revealed by the ninhydrin-carbon dioxide determination at pH 2·5 (van Slyke, Dillon, MacFadyen & Hamilton 1941). (Found: carboxyl N, 7·8, 8·0, 7·7. $C_5H_{12}N_2O_3S$ requires N, 7·8 %.)

Degradation to α-amino-n-butyric acid

Hydrogenolysis of the crystalline substance by means of Raney nickel (compare Fonken & Mozingo 1947) gives α-amino-*n*-butyric acid.

A solution of crystalline toxic factor (30 mg.) in water (10 ml.) was heated for 45 min. on the steam bath with Raney nickel (about 1·0 g.). The solution was then

filtered free from Raney nickel and the nickel washed with hot water (50 ml.). The combined filtrate and washings were evaporated to dryness under reduced pressure on the steam bath and the dry residue recrystallized from a small volume of aqueous ethanol giving colourless needles (13 mg. yield 79 %), m.p. 270° C (decomp.) alone and m.p. 276° C mixed with an authentic specimen of α-amino n-butyric acid (m.p. 276° C) obtained by the hydrogenolysis of methionine (Found: N, 13·5; calculated for $C_4H_9NO_2$, N, 13·6 %). The product, further, gives a single spot on paper chromatograms running parallel with authentic α-amino-n-butyric acid in phenol (R_F 0·7) and acetone (R_F 0·55).

Acid hydrolysis

Hydrolysis of crystalline toxic factor with hydrochloric acid (6 N) for 24 hr. in a sealed evacuated tube at 110° C results in almost complete degradation of the original substance. Paper chromatography of the hydrolysate in phenol reveals at least four distinct ninhydrin-reacting spots, one of which is probably due to homocysteic acid but none of which has been identified unequivocably. The production of a number of new ninhydrin-reacting substances on acid hydrolysates at first led us to think that the substance was a peptide. We could not, however, reconcile the analytical data with a peptide structure and, in addition, it was subsequently observed that DL-methionine sulphoxide when acid hydrolyzed under similar conditions gives on paper chromatograms a pattern of new ninhydrin-reacting spots.

DISCUSSION

The interaction of nitrogen trichloride and proteins

The reaction between 'agene' and a finely divided protein is a complex process, and the optimum conditions for our own particular needs were largely determined by a process of trial and error. Of the proteins investigated the behaviour of zein has been most closely studied. It has the practical advantage of being readily available commercially in a suitable form, although, as described in the experimental section, the commercial powder requires to be ground to a finer state of subdivision to ensure maximum toxicity from a given weight of protein. In our first experiments commercial zein was purified by precipitation from a solution in 70 % ethanol, the glutinous precipitate being then comminuted by prolonged grinding under acetone. For large-scale operations such a process was impracticable and was, indeed found subsequently to be unnecessary.

The amount of 'agene' (calculated as nitrogen trichloride) required for the saturation of zein varies with the particle size and with the moisture content of the protein. With finely powdered material at a moisture content of 7 to 10 % the saturation value is approximately 20 mg. NCl_3 per g. zein, giving a product toxic at the level of 4 g. in dogs and 0·8 g. in rabbits. At 16 % moisture content the saturation value is considerably higher (40 mg. NCl_3 per g. zein), but the product is non-toxic. The toxicity of zein agenized at the level of 20 mg. NCl_3 per g. protein is destroyed if the product is 'agenized' further at a higher moisture content.

Zein 'agenized' at different levels also varies considerably in the ease of digestion by pancreatin. The best compromise between yield of toxicity and ease of digestion and dialysis is given by zein treated at the level of 5·5 mg. NCl₃. This treatment gives a product toxic at the level of 25 g. in dogs and 5 g. in rabbits. After digestion for 3 days at 37° C followed by 7 days' dialysis the undigested solid residue is non-toxic and all toxicity is recovered in the soluble digest.

With zein saturated with 'agene' at 20 mg. NCl₃ per g. protein, about 60 % of the protein resists digestion with pancreatin, the undigested residue is toxic and no toxicity is recovered in the soluble digest. Partial acid hydrolysis of the undigested residue with 10 N-hydrochloric acid for 10 days at 37° C yields soluble material toxic at the level of 0·9 g. in rabbits.

These observations illustrate the complexity of the reaction with an intact protein. Particle size of the protein is important, and the moisture content is critical for the formation of the toxic factor. It is to be noted that under certain conditions treatment of the protein with an excess of NCl₃ results in the destruction of toxic material already formed.

Characterization of the toxic factor

The identification of the pure toxic factor as an amino-acid containing sulphur and five carbon atoms and yielding α-amino-*n*-butyric acid on hydrogenolysis implies that the essential precursor in the protein molecule is a methionine residue. Reiner and co-workers (private communication) have provided additional evidence for methionine as the precursor by deriving methionine sulphone from the toxic factor by treatment with hydrogen peroxide.

Under optimum conditions of treatment with NCl₃ only a small proportion of the methionine residues present in the protein form the toxic substance. Thus with denatured egg albumen the absorption of NCl₃ is sufficient to interact with twice the amount of methionine present but only about one-thirtieth of the methionine appears as toxic factor. It is impossible to say how far this may be due to some other specific grouping linked with methionine in the protein being essential, because of the complicating factors, already mentioned, of the reaction with NCl₃. That some specific structural or stereochemical configuration of the methionine residues in the intact protein is necessary for the production of the toxic factor is suggested by the fact that it is not produced when NCl₃ reacts with free methionine. Methionine has been treated with NCl₃ under a variety of conditions: supported on kieselguhr at various moisture contents, supported on starch, in a free-powdered form, suspended in carbon tetrachloride and dissolved in aqueous solution. Under none of these conditions has there been any observable production of the toxic substance. The products have invariably been methionine sulphoxide and sulphone.

The molecule $C_5H_{12}N_2O_3S$ is derived from methionine, $C_5H_{11}NO_2S$, by the addition of one atom each of nitrogen, oxygen and hydrogen:

$$\text{HOOCCHCH}_2\text{CH}_2\underset{\underset{\text{CH}_3}{}}{\overset{\overset{\text{O}}{}}{\text{S}}}\text{NH}$$
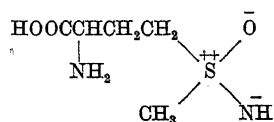
The foregoing evidence suggests strongly that these extra atoms are in fact attached to the sulphur atom, and the most probable formulation of the molecule is then (I) which is presumably stabilized by resonance in the same manner as a sulphone. Sulphone analogues of this type have not previously been described, but a synthesis of the toxic factor by the action of hydrazoic acid on methionine sulphoxide (Bentley, McDermott & Whitehead 1950b) confirms the structure (I) unambiguously.

This molecule possesses two asymmetric centres, and the toxic factor isolated from 'agenized' zein will therefore be a mixture of two diastereoisomerides, assuming that the factor originates in an L-methionine residue and that racemization does not occur at the α-carbon atom during the isolation procedure. It remains of interest therefore to separate the two isomers and to compare their toxicities. We have reported a preliminary experiment along these lines with material isolated from agenized zein (Bentley et al. 1950a) using picric acid (Lavine 1947), but with the greater quantities of the toxic factor now available synthetically it is hoped to effect a complete separation.

[Note added in proof. The two fractions A (m.p. 248° C) and B (m.p. 234° C) obtained from the crystalline factor X (m.p. 234° C) by treatment with picric acid have the same R_F values on the paper chromatogram and the elementary analyses correspond closely (Bentley et al. 1950a). Both fractions are toxic to rabbits but A appears to be more potent than B. Since this paper was submitted Dr H. W. Thompson, F.R.S., has kindly determined the infra-red absorption spectra of X, A, B and of synthetic toxic factor (m.p. 236° C) prepared by the action of hydrazoic acid on 1-methionine sulphoxide (Bentley, McDermott & Whitehead 1950b). The spectra of A and the synthetic material are identical, but the spectra of X and B differ from those of A and the synthetic material in several spectral regions. This difference appears to be greater than can be accounted for on the basis of stereoisomerism and it would therefore appear that the separation effected with picric acid cannot be considered solely as a partial resolution of diastereoisomers.]

REFERENCES

Albanese, A. A. 1940 J. Biol. Chem. 134, 467.
Bentley, H. R., Booth, R. G., Greer, E. N., Heathcote, J. G., Hutchinson, J. B. & Moran, T. 1948 Nature, 161, 126.
Bentley, H. R., McDermott, E. E., Pace, J., Whitehead, J. K. & Moran, T. 1949a Nature, 163, 675.
Bentley, H. R., McDermott, E. E., Pace, J., Whitehead, J. K. & Moran, T. 1949b Nature, 164, 438.

Bentley, H. R., McDermott, E. E., Pace, J., Whitehead, J. K. & Moran, T. 1950a *Nature*, 165, 150.

Bentley, H. R., McDermott, E. E. & Whitehead, J. K. 1950b *Nature* (in the Press).

Bentley, H. R. & Whitehead, J. K. 1950 *Biochem. J.* 46, 341.

Callan, T. P. & Toennies, G. 1941 *Industr. Engng. Chem.* (Anal. ed.), 13, 450.

Campbell, P. N., Work, T. S. & Mellanby, E. 1950 *Nature*, 165, 345.

Cole, S. W. 1923 *Practical physiological chemistry*, 9th ed. p. 93. Cambridge University Press.

Consden, R., Gordon, A. H. & Martin, A. J. P. 1944 *Biochem. J.* 38, 224.

Dent, C. E. 1948 *Biochem. J.* 43, 169.

Fonken, G. S. & Mozingo, R. 1947 *J. Amer. Chem. Soc.* 69, 1212.

Heathcote, J. G. 1949 *Nature*, 164, 439.

Isherwood, F. A. 1946 *Biochem. J.* 40, 688.

Lavine, T. F. 1947 *J. Biol. Chem.* 169, 477.

Mellanby, E. 1946 *Brit. Med. J.* 2, 885.

Mellanby, E. 1947 *Brit. Med. J.* 2, 288.

Moran, T. 1947 *Lancet*, 2, 289.

Newell, G. W., Erickson, T. C., Gilson, W. E., Gershoff, S. N. & Elvehjem, C. A. 1947 *J. Amer. Med. Ass.* 135, 760.

Olley, J. 1945 *Biochem. J.* (Proc.), 39, 1.

Radomski, J. L., Woodward, G. & Lehman, A. J. 1948 *J. Nutrit.* 36, 15.

Randall, S. S. & Martin, A. J. P. 1949 *Biochem. J.* (Proc.), 44, xi.

Reiner, L., Misani, F. & Weiss, P. 1950 *Arch. Biochem.* 25, 447.

Stein, W. H. & Moore, S. 1948 *J. Biol. Chem.* 176, 340.

Van Slyke, D. D., Dillon, R. T., MacFadyen, D. A. & Hamilton, P. 1941 *J. Biol. Chem.* 141, 627.

The Bell Telephone Laboratories—an example of an institute of creative technology

By Mervin J. Kelly

Bell Telephone Laboratories, New York, N.Y.

[Plates 21 to 23]

Introduction

The part that science and technology have played in influencing the economic, social and political patterns of western society and in enriching the lives of its people has steadily increased during the last century. The scope and character of this influence have varied widely from country to country. Traditions, mores, maturity, size, patterns of education, and many other factors have been elements in bringing about the variations.

The influence has probably been most profound in the United States, principally, I believe, because of its youth, size, and patterns of education. Beginning some four or five decades ago, my country has been transformed at an increasingly rapid tempo from primarily an agricultural society to predominantly an industrial one under the driving force of an expanding body of science and technology. So completely have they dominated the pattern of our growth that when the man in the street speaks of 'progress', he usually means scientific and technological progress.

Until the beginning of this century, the work in applying new scientific knowledge to new facilities and instrumentalities for society was quite a hit and miss process. The inventor, having little or no direct contact with pure science, took the first steps whose end product was the inventor's model. Then the engineer, who at that time was largely a graduate from the drafting board or from the shop, reduced the inventor's model to the design of a new product for manufacture. While progress was made, the procedure was slow, inefficient, and the intervals of time were quite long between the availability of new scientific knowledge and the appearance of new products made possible by it.

The break from this pattern began at the turn of the century with the appearance in industry of men trained in the scientific method of research. They were the pioneers of industrial research and established industry's first laboratories of applied science. Dr Jewett of the Bell System and Dr Whitney of the General Electric Company were among the first. Other men with similar training followed them, initially at a slow rate, but, as they proved their value, the number rapidly grew. We now have some 2500 industrial research laboratories. They vary in size from the little fellow with less than fifty employees to the giant with more than a thousand. As they have grown in size, aged, and learned through experience, their quality has steadily improved.

To keep pace with the evolution of its research laboratory and take advantage of the opportunities accruing from the adoption of the scientist and his methods, the engineering organization of industry has undergone major change. Its relatively simple operation, in the last century, of transforming the inventor's model into a design for manufacture, performed largely by empirical methods, has now expanded into many successive interlaced operations. Each, as it has matured, employs more of the scientific method and of fundamental analysis in the solution of its problems.

The applied science laboratory and these engineering functions have been integrated in various ways into the organization of our industrial corporations. Whatever the organization may be, the work is a continuously flowing scientific and technical operation that begins at the forefront of the areas of pure science that are significant to the industry, and proceeds through successive steps of research, development, design, and engineering for manufacture of the new product or facility.

There has been so much emphasis on industrial research and mass-production methods in my country, that even our well-informed public is not sufficiently aware of the necessary and most important chain of events that lies between the initial step of basic research and the terminal operation of manufacture. In order to stress the continuity of procedures from research to engineering of product into manufacture and to emphasize their real unity, I speak of them as the single entity 'organized creative technology'.

HISTORICAL

I am using the Bell Telephone Laboratories and its operations as an exemplification of this unity. But first I must make sure that all of you are familiar with us.

The Laboratories is an organization of about 5700 scientists, engineers, and supporting staff—2200 are scientists and engineers. It serves the Bell System—an operating or service organization which provides nation-wide telephone service for the United States. Some 34 million telephones are owned and operated by the Bell Companies and are interconnected with some 7 million additional telephones operated by a large number of relatively small independent companies.

The Laboratories provides the Bell System with new communications technology and the information for its implementation with new systems and facilities. Its work begins at the forefront of applied science, extends across the entire scope of creative technology through research, development, design, and engineering. The end products of its programmes are the designs and specifications for manufacture of the technical facilities that it has created and the engineering practices to be followed in their operation and maintenance.

Since the earliest days of the telephone, the Bell System has been responsible for its own technology and for the manufacture of the major portion of its technical facilities. The Western Electric Company, a subsidiary, has always been the manufacturer.

The Bell System is unique among organizations providing service in its responsibility for its own technology and for the manufacture of its technical facilities. The service of telephony differs fundamentally from that of others, such as power, gas

and water, in its technical complexity, its essential national integration, and in its two-way character. The telephone system of the United States can be viewed as a single, integrated, highly technical machine in which electrical currents that are very small and complex in wave form are sent from any one of more than 40 million points to any one of all the others. The intensity (amplitude) and wave form of these currents must be accurately controlled and, because of their extremely low value, interfering electrical currents of a great variety of kinds must be either excluded from circulating in the network or maintained at a level well beneath that of the 'intelligence' currents.

This integration of technology, manufacture, and operation has permitted the rapid application of new scientific knowledge to new and improved telephone services and the realization of sound engineering and economic solutions to the large variety of the complex problems of a nation-wide telephone service. This has been a major element in providing the American public with the most complete telephone service of the world and at a cost, in terms of the income of its people, so low that we make use of telephony in our daily economic and social life to a greater extent than any other people.

It was not until 1925 that all the functions of creative technology were consolidated into a single organization. At that time the Bell Telephone Laboratories was formed as a subsidiary corporation. While corporate and organizational unification occurred only 25 years ago, our operations in creative technology had been under development during the preceding 20 years. Dr Jewett, the first president of our Laboratories and one of the first scientists engaged in Bell System work in technology, exerted an increasing influence on all sectors of our creative technology independent of its organizational pattern.

We have, therefore, a background of almost a half-century of experience and a professional staff whose age distribution in the years from 25 to 65 is now substantially uniform. Through trial and error with continuous appraisal of different ways for organization of work, types of men and their training, and of housing and facilities, our present patterns of work, of men and their training, and our housing and facilities have evolved.

In the half-century in which our Laboratories has evolved, the physical sciences underlying telecommunications have made greater progress in advancing knowledge of the laws of nature and the structure of matter than in any equal period of history. The increase in fundamental knowledge has been at a rate so rapid that we have been faced with an increasingly complex problem in maintaining a close linkage between the forefront of our applied research and that of pure science.

It was during this same period that the graduate schools of our country entered into full participation in research in the physical sciences as an equal partner with the older society of Europe. At the beginning of the century there were only a few of our universities in which research at the forefront was in progress.

From this small beginning there has been an increasingly rapid expansion in the volume of research and in the number of universities and institutes of technology participating, until to-day there are active and effective programmes in research in the physical sciences in almost a hundred universities and institutes of technology.

28-2

This expansion has been accompanied by the training in each year of a greater number of men at the graduate level in the physical sciences and in the methods of scientific research.

Our Laboratories, therefore, grew in size and matured in the scope and character of its work during the period of rapid expansion in research in the physical sciences and in the number of men capable of working in the most modern methods of scientific research. This presented us with a most difficult problem of maintaining a leadership and a professional staff of adequate competence and training to establish programmes based on the most recent scientific knowledge and to carry them out with methods and instrumentation that kept pace with the progress of pure research.

The organization of our work

To present the Laboratories as a mature institute of creative technology, I shall consider its operations under three general headings.

The first includes all of the research and fundamental development. This is our non-scheduled area of work. It provides the reservoir of completely new knowledge, principles, materials, methods, and art that are essential for the development of new communications systems and facilities. Arising in the course of the work and forming an important part of it is a large fraction of those creative concepts of our Laboratories that constitutes invention.

Its programmes extend across all sectors of science that may contribute to the advancement of the communication art. They are carried out in sufficient volume that there is the minimum reasonable time lag between an advance in pure science and our realization of its contribution to our reservoir of new knowledge.

The second we call 'systems engineering'. Its major responsibility is the determination of the new specific systems and facilities development projects—their operational and economic objectives and the broad technical plan to be followed. 'Systems engineering' controls and guides the use of the new knowledge obtained from the research and fundamental development programmes in the creation of new telephone services and the improvement and lowering of cost of services already established. In determining the new development projects, 'systems engineering' considers the content of the reservoir of new knowledge awaiting application and the opportunities for its use in the interest of the telephone user. The projects that are activated are those that give greatest promise for user benefit. It attempts to insure that the technical objectives of the development projects undertaken can be realized within the framework of the new knowledge available in the reservoir and present engineering practice.

The third encompasses all specific development and design of new systems and facilities. The work is most carefully programmed in conformity with the plan established by the systems engineering studies. Our research and fundamental development programmes supply the new knowledge required in meeting the objectives of the new specific developments.

Before a new development project is activated, 'systems engineering' establishes well-considered performance and cost objectives. It determines the general technical

line of attack sufficiently well to insure that there is adequate fundamental knowledge available in our reservoir for its realization. The development project is carried out under the charter established by systems engineering. There are three well-ordered and successive steps in its performance. The end products are the specification of designs and materials for the manufacture of the system or facility, and the engineering practices to be followed in its operation and maintenance in service.

I shall now turn to a more detailed examination of the programmes of each of the three areas, their interrelationships, and the types and training of the men that we have found best for the work.

RESEARCH AND FUNDAMENTAL DEVELOPMENT

Researches in the sectors of physics, chemistry, and mathematics that give promise of contribution to advances in telephony are carried out in the forefront of the first area. Solid state physics, including magnetism, piezo-electricity, dielectrics, and semiconductors; physical electronics; electron dynamics; acoustics; electromagnetics; mathematics; organic and physical chemistry of synthetic plastics and rubber; corrosion chemistry; physical metallurgy; and fundamental mechanics are a typical but not comprehensive list of subjects.

The research area provides the coupling between the ever-advancing forefront of pure science and the forward march of our communications technology. The closer the coupling, the more completely will we keep our advancing technology in step with the progress of science.

Years of experience have taught us that the type and quality of men selected for our research, the environment that we provide, and the distance in their work that we ask them to penetrate beyond the forefront of creative technology are the most important factors in determining the closeness of coupling—the effectiveness.

Inspired and productive research in industry requires men of the same high quality as is required for distinguished pure research in our universities. We select our young men of research from among the most able and promising of the doctorate and post-doctorate of philosophy students of our graduate schools. We have no difficulty in attracting men of this type and quality.

They must be given freedoms that are equivalent to those of the research man in the university. This is, indeed, difficult in industry, but we are approaching that ideal. We give much attention to the maintenance of an atmosphere of freedom and an environment stimulating to scholarship and scientific research interest.

Our research staff make a practice of publishing the new scientific knowledge resulting from their programmes in the same manner as do the research men of the universities. In 1949 the members of our research staff published more than 200 scientific papers resulting from their research programmes. There are now seven treatises in preparation that will appear in book form in the Bell Telephone Laboratories-Van Nostrand series. Our scientists attend and participate in the meetings of their professional societies. They have a most active science seminar with programmes extending over a wide area of scientific interest. Fourteen

British scientists have been among the visiting speakers at the seminar during the past four years.

It is most important for the scientists to confine their efforts to the area of research. If they extend the area of their effort even to that of fundamental development (the area of work that immediately follows research) they tend to lose contact with the forefront of their field of scientific interest. In time, a considerable fraction will lose their productivity in research.

To make possible and to encourage this concentration of the attention of the men of science to research, we provide, through organization and stimulated association, an intimate tie between research and fundamental development, the next step in the chain of events from research to manufacture and use. In this way the programmes of research are taken over at a well-considered point by fundamental development, where they are extended and enlarged upon to supply the body of basic technology essential for the specific development and design of systems and facilities. While the fundamental development work is done in the best research tradition, it has a large content of the technologic, and economic considerations begin to be a factor in its programmes.

Staff Members for fundamental development are drawn from our research groups by selecting those having technologic and engineering aptitudes and interests who prefer to move into development and by recruitment from among the most promising of the graduate students of our schools of applied science, such as Massachusetts and California Institutes of Technology.

A typical example of a programme of this area will be helpful in making clear the distinctions between research and fundamental development and their inter-relationships. A few years ago members of our solid state physics research group focused their attention on the mechanisms of conduction in semiconductors, such as germanium and silicon. Important new knowledge of the conduction process acquired in the researches served as an inspiration and background for experiments with means and techniques for controlling the flow of the current in a semiconductor.

Out of this work came solid state amplifiers with characteristics quite similar to those of the thermionic vacuum tube amplifier. This is a discovery of major import-ance to communication technology. In mid-1948, after the scientists had rounded out their understanding of the physics of the amplification process, the Laboratories made public announcement of the solid state amplifier, which we called the 'transistor'. Concurrently, the scientists working on the project published the results of their researches, and have since published new findings and frankly discussed this new phenomenon with their peers.

In accord with our policy of concentrating the efforts of our scientists on research, we immediately formed a closely associated fundamental development group to acquire that body of technological knowledge essential to the development and design of transistors for the many specific communications applications that would certainly follow. They have interested themselves in such problems as the factors controlling the bandwidth of amplification; the noise figure; the amount of amplification possible per stage; energy levels of output; basic materials, processing and structure studies essential for controlled development and design of transistors for specific functions;

etc. Their work has been followed with sympathy and interest by the scientists in solid-state research. The fundamental development group has continuously consulted with the men of research on many of their problems. The contact of the scientists with the fundamental development work, as is to be expected, has not only been of great value to development but has stimulated further research work.

This pattern repeats itself again and again. A research programme is initiated. Then as new knowledge that gives promise of worth-while application is obtained, a fundamental development team is activated. It builds a background of basic technology under the watchful eye and with the consultative aid of the men of research.

Approximately 30 % of our professional staff work in the area of research and fundamental development. This provides adequate effort to maintain close coupling between our technology and science, and supplies a sufficient volume of new knowledge to feed into the programmes of our systems and facilities development area. That is to say, this division of effort gives good balance in our over-all programme.

Systems engineering

One of the principal responsibilities of systems engineering is technical planning and control. In the planning an appraisal is made of the various technical paths that can be followed in employing the new knowledge obtained by research and fundamental development in the specific development and design of new systems and facilities. The determination of the most effective use of new knowledge in the interest of the telephone user is the guiding principle of the planning studies. The most effective use may be the creation of new services, the improvement of the quality of existing services, the lowering of their cost, or some combination of these three. As the technology of communication has broadened and become more complex, the choice of the technical paths to be pursued in the instrumentation of the new technology has become increasingly difficult. It is this situation that has led to the evolution of the systems engineering function as a mechanism of guidance and control.

Systems engineering has intimate knowledge of the telephone plant and its operation and maintains close contact with the engineers of the operating organizations. The teamwork of operating engineers and our systems engineers makes available to the Laboratories in a most effective way the knowledge of the telephone system's needs and the opportunities for economy and improvement.

Systems engineering also maintains close association with our research and fundamental development work. It knows intimately the content of our new knowledge reservoir. It has equally effective liaison with the men and their work in the specific systems and facilities development area.

It integrates the knowledge from operations, research and fundamental development, and specific systems and facilities development. With this as a background, it makes exhaustive studies that appraise and programme development projects for new systems and facilities. Each study outlines the broad technical plan for a development, its objectives, and its economic and service worth. In many of

the studies it is recommended that no development be undertaken at the time. Administrative action is then taken with these studies as a guide to the selection and activation of the projects of our specific development programmes.

As the development organization proceeds with a project, systems engineering maintains close contact, continuously appraises the results, and amends the objectives and plans as required. Service trials are generally needed during the course of development. It organizes the trials in co-operation with operating engineers and participates in the tests and the evaluation of results. When the system is standardized and placed in manufacture, systems engineering follows service performance of first installations and coordinates the 'growing pains' engineering that is ever with us on new systems as they enter service. It finally participates in the evaluation of the service and economic worth that are experienced.

Typical examples of recent systems engineering studies that have led to development and standardization are: television transmission over coaxial cables, a broadband microwave radio repeatered communication system, an automatic message accounting system, a mobile radio subscriber telephone system, and a new subscriber telephone set.

Systems engineering has another important responsibility. It recommends the levels of the various technical standards that are important elements in determining the quality and reliability of telephone service. Frequency bandwidth for voice transmission, noise levels on circuits, distortion and range of energy level of the speech currents, cross-talk levels, standards of protection against man-made and nature's interference with service are typical.

With advancing technology the levels of these various standards have gradually been raised. This has been reflected in improvements in quality, speed, and reliability of service. There must be a balance between service costs and rigour of these standards. It is systems engineering's responsibility to correlate with other organizations of the Bell System the factors involved in keeping standards and costs in balance so that a well-considered portion of economies available through advances in technology are used in raising quality and reliability and in increasing the speed of service.

Approximately 10 % of our scientific and technical staff are allotted to systems engineering. Its staff members must supply a proper blending of competence and background in each of the three areas that it contacts: research and fundamental development, specific systems and facilities development, and operations. It is, therefore, largely made up of men drawn from these areas who have exhibited unusual talents in analysis and the objectivity so essential to their appraisal responsibility. For some of the more complex studies, the staff is augmented by borrowings of men having special knowledge or abilities from one or more of the three areas it coordinates.

Specific systems and facilities development

The work of specific systems and facilities development is closely programmed. Its projects are organized in the patterns that the studies of systems engineering prescribe.
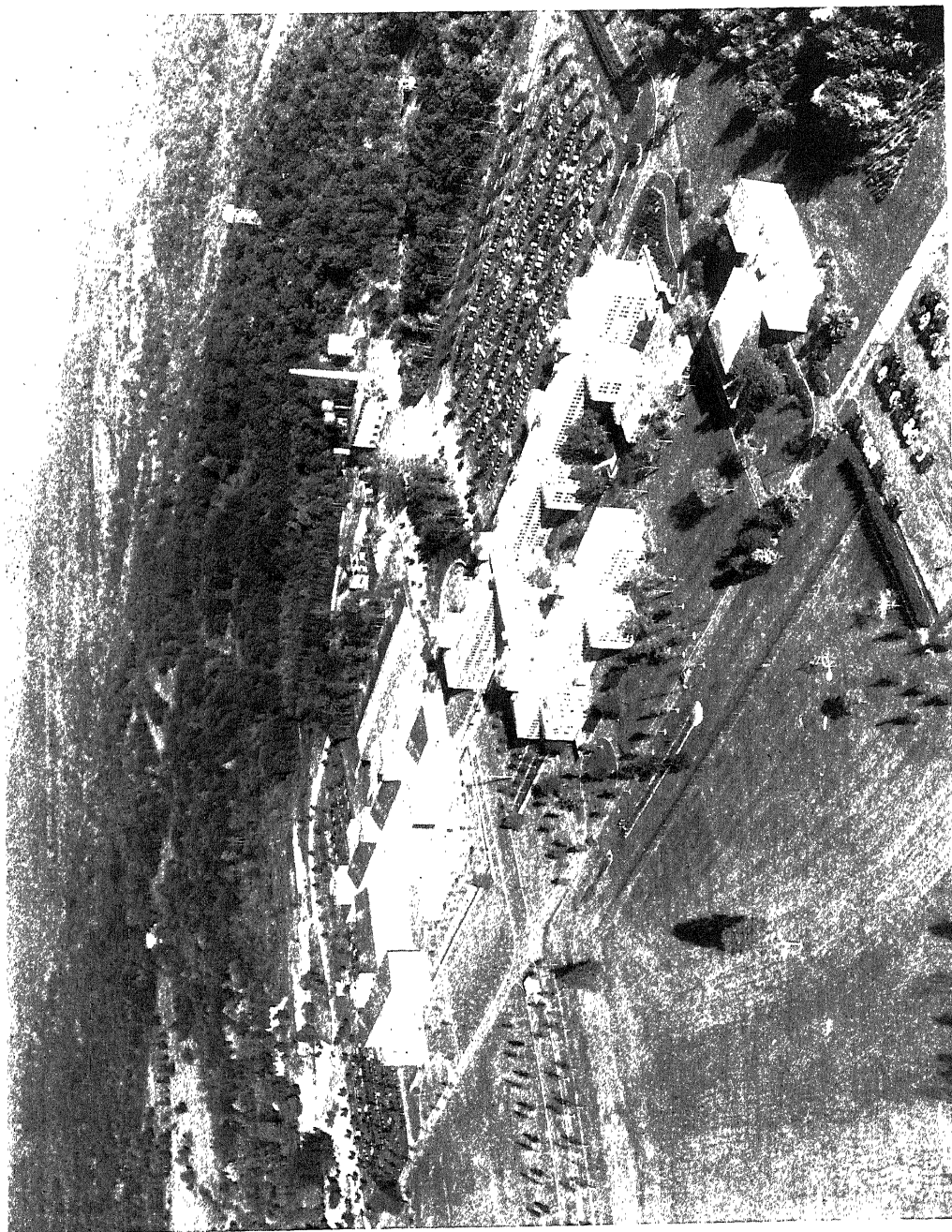
FIGURE 1. Aerial view of Murray Hill Unit of Bell Telephone Laboratories.

FIGURE 4. A typical physical research laboratory.

FIGURE 3. Placement of a panel of the removable metal partition.
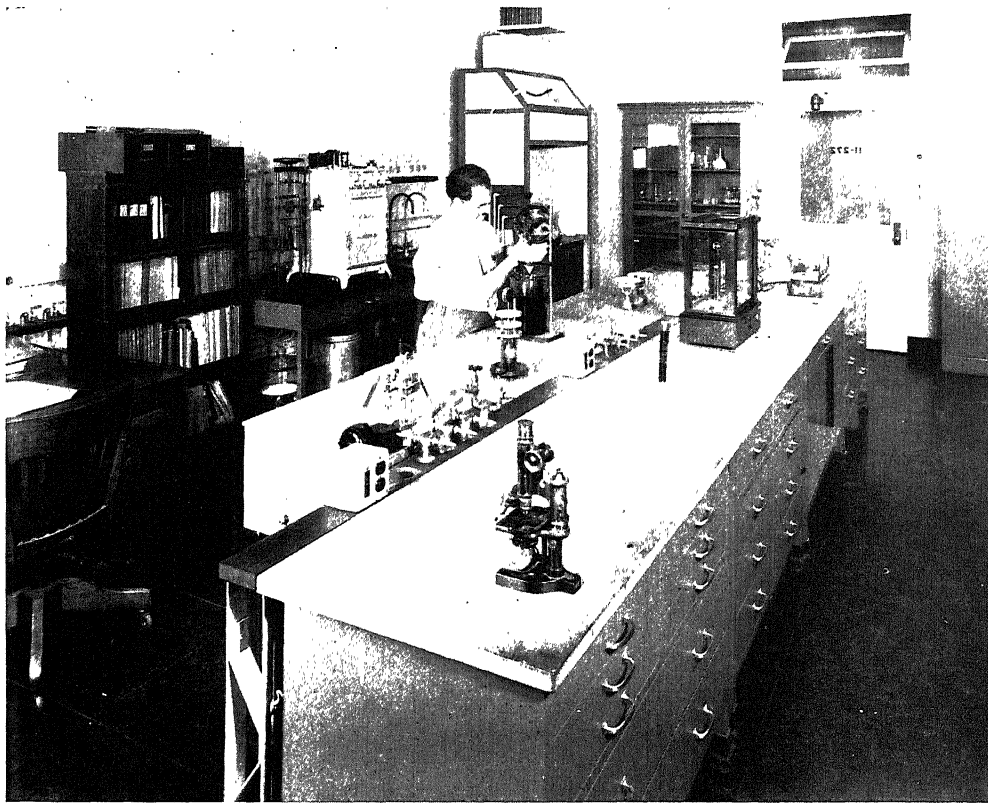
FIGURE 5. A typical chemical research laboratory.
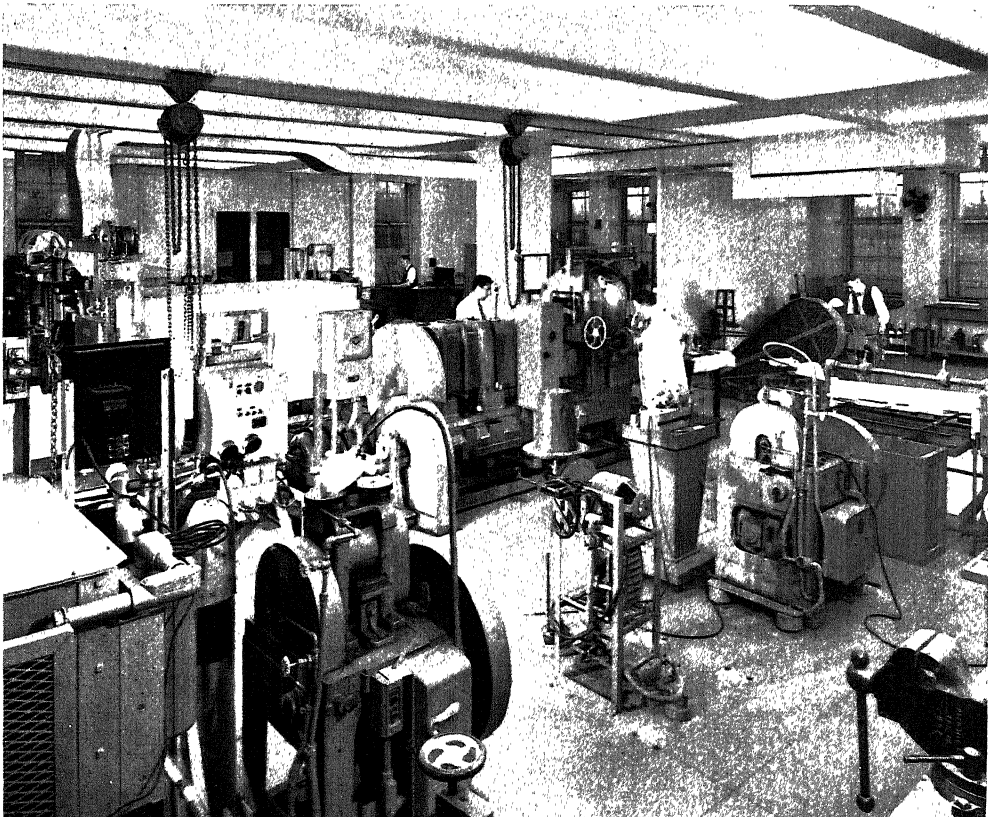
FIGURE 6. The metal fabrication unit of the metallurgical research laboratory.

The development, while a continuous operation, is done in three distinct stages. In the first, a laboratory model is evolved that meets the functional requirements of the systems engineering study. These are not always the same as initially established, for new information obtained during the course of development may well modify them. Systems engineering participates in their modification. Development is carried out within the framework of the new knowledge available in our new knowledge reservoir and of already established development and engineering practice. The model is thoroughly tested and proven for functional performance. Most intimate working contact is maintained between the specific development group and those of fundamental development throughout this stage.

In the next stage the design for manufacture and use is created. While retaining the functional performance of the laboratory model, it also meets the requirements of manufacture at lowest cost consistent with providing the specified service at lowest complete service cost. (The capital and depreciation charges that are directly related to manufacturing cost comprise only a portion of the complete cost of service.) In achieving these economic objectives in the design, the development group is the focal point of a closely integrated trifurcated team effort with manufacturing engineering of Western Electric as interpreters of, and important contributors to, design form for lowest manufacturing cost and with systems engineering as interpreters of operating requirements that are essential to lowest complete service cost.

The end product of stage two is a pre-production model. Here we are approaching final design. Western Electric produces a number varying from a few to a few hundred in their job shops and with preliminary tools. These models are placed in service under the close observation of the systems engineering and specific development groups of the Laboratories and the appropriate engineering groups of the operating companies.

The findings of the service tests and Western Electric's experience in producing the models are used as background for our final freezing of design and Western Electric's planning and tooling for quantity production. Design drawings and specifications are then prepared for manufacture, and engineering practices for the technical operations involved in supplying telephone service.

Some 60% of the members of our professional staff are engaged in the development and design of systems and facilities. They are principally graduate engineers—electrical, mechanical, chemical, and metallurgical—from our technical schools and universities. The procedures of development and design have become increasingly fundamental and analytical. At the beginning of the century the work was largely empirical and done principally by graduates from the shop and drafting board.

We believe the maturity of science and fundamental technology that has been achieved offers opportunity for an even more fundamental and analytical approach in these final steps of creative technology. After much consideration of this problem during the post-war years, we concluded that a more fundamental approach can best be promoted if some of the engineers of specific development and design receive more fundamental training than they now obtain in the four- and five-year engineering curricula. We have, therefore, established within the Laboratories

a training course at the graduate school level in which selected engineering recruits receive added training in science and the technology of communications while on the job.

This completes the description of our work patterns and types of men involved in the successive steps of creative technology comprising research, fundamental development, systems engineering, and specific development and design. While programmes are systematically organized in this pattern and each member of the professional staff is assigned and works in that sector of creative technology for which his training, experience, and natural aptitudes best fit him, there exist important formally organized and informal team efforts within and across the boundary of each area.

So great is the complexity of applied science and technology of telecommunications that much specialization is required. The leaders of our programmes give a considerable portion of their time to the development of formal and informal co-operation. For example: While there are men with the necessary mathematical training and facility for the analytical work that normally is encountered in each of the areas of our work, the mathematicians of our research area devote at least one-fourth of their time to consultation and to aiding the analysts of the different areas of development in the solution of their problems. Such co-operation is informal and initiated by the men of development requiring the help.

Through the years our research and development leaders have developed patterns of informal co-operation and the habit of going to the expert, whether he be a mathematician, a metallurgist, an organic chemist, an electromagnetic propagation physicist, or an electron device specialist. This has made it possible to focus the full power of our organization on a particular problem. The organization's capacity for the solution of telecommunication problems is much greater than the sum of the capabilities of the individuals. Teamwork and co-operation are the important elements that bring about this amplification of strength.

The time that management gives to developing an environment favourable to teamwork and co-operation is most rewarding. It not only assures a high level of effectiveness in current programmes but, perhaps even more importantly, it fosters a continuing development of the members of staff and the institute as a living organism. It provides an organization of the flexibility and resilience so essential in meeting the challenges of a rapidly changing world of science and technology.

LABORATORY HOUSING, FACILITIES, AND SERVICES

Laboratory housing, facilities, and services is a topic that follows closely men, methods, and organization in its importance in the building of an efficient institute of creative technology. By the mid-1930's we fully recognized the limitation to our effectiveness imposed by the non-functional housing and old designs of facilities and services. A small committee was formed to determine the design of housing, facilities, and services best suited to effective laboratory operations.

In the later phases of the study the Bell System's architects were called in, and in co-operation with them we designed a functional structure to accommodate

1000 employees and with facilities and services complementary to the design. This laboratory was built on a 300-acre tract at Murray Hill, New Jersey, some 20 miles from New York City. It was occupied in 1941 and made a large contribution to the effectiveness of our war programmes. At the close of hostilities we made those modifications in design that experience indicated desirable and began the construction of an adjacent unit to house some 1600 employees. Its occupancy will be completed this year. The two units are connected by a passage-way at the second-floor level. An aerial view of the two buildings is shown in figure 1, plate 21.

The motif of the functional design of the buildings has been complete flexibility of room size with easily removable partitions for all interior walls. An institute of creative technology must be provided not only with space for laboratories and shops but also offices, conference rooms, drafting rooms, and clerical work. The structures provide two types of functional space, one for laboratories and shops and the other for offices, etc.

Laboratories and shop space comprise the main stems of the buildings. The stems are long, rectangular blocks of different lengths but of a uniform width of 54 ft. Our study indicated that we could most effectively meet the need for rooms of different sizes by having a repetitive 6 ft. module along the long axis of the rectangle. A corridor, 7 ft. wide, runs the length of the long axis of the rectangle. One side of the corridor is on the axis of the rectangle, thus providing rooms on either side of the corridor differing 7 ft. in depth. On one side of the corridor the depth is 27 ft. and 20 ft. on the other. This makes possible rooms 27 ft. in depth with widths that may be any integral multiple of 6 ft. on one side of the corridor and rooms 20 ft. in depth with widths that may be any integral multiple of 6 ft. on the other.

The repetitive 6 ft. unit was designed to include in its outer wall all of the facilities that might be needed in any laboratory space. Each 6 ft. unit was provided with a window for natural lighting and ventilation, artificial lighting fixtures, heating, ceiling inserts for support of apparatus, telephone, and access to fourteen types of services comprising various gases under pressure, electrical supply of different characteristics, as well as other services normally involved in telecommunications research and development.

Figure 2 is a diagram of two of the standard laboratory modules of 6 ft. and the distribution of service. This makes evident the manner in which the outside wall is used as the source of all services.

The adoption of the 6 ft. module with all services in the outer walls was predicated upon the use of a versatile system of wainscoting and partitioning to allow quick—often week-end—changes of work space. The entire building has a metal lining which is readily removable, and through which free access is given to the mechanical services.

The removable metal partition panels of the subdividing walls are the exact height to extend from the floor level to the ceiling beams, and are some 4 ft. wide. Each panel is 3 in. thick and consists of two layers of metal separated by mineral wool to make the panels fire-resistant and satisfactory in the elimination of sound transmission. In order to accommodate shelves, blackboards, or service piping supports, a keyholed device may be inserted between any two panels on 6 in.

vertical centres. The corridor walls are also made of the same subdivision panels, which are interchangeable with panels equipped with doors and transoms where these are needed.

Figure 3, plate 22, shows the placement of one of the removable metal partition panels in the construction of a subdividing wall. The metal lining of the outer wall is removed below the window level and the distribution piping for gases under pressure can be seen.
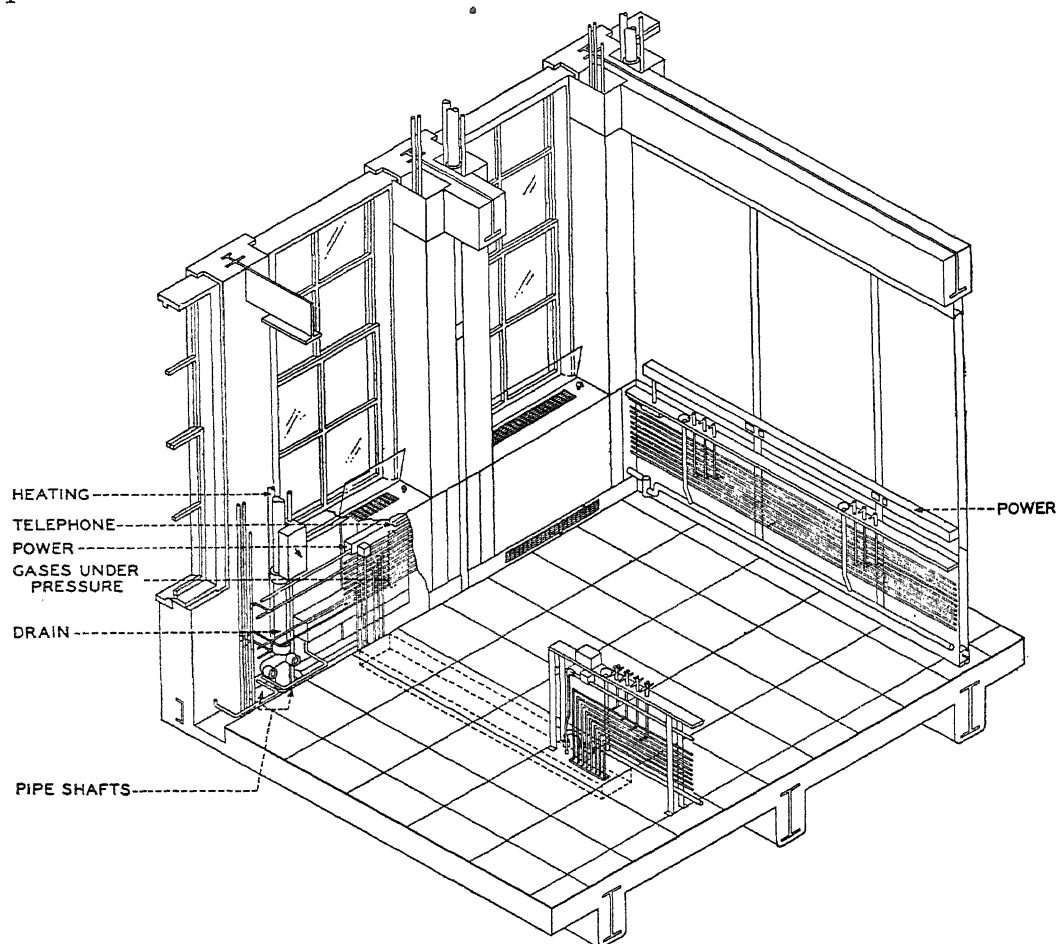
FIGURE 2. Diagram of two of the 6 ft. laboratory modules.

Laboratory furniture and facilities have been designed to complement the functional structure of the building. Every item of furniture and facilities, including even the chemical fume hood, has been made completely detachable and readily transportable. With the removable partitions and the furniture and facilities of special design, laboratory rooms of any required size can be readily constructed and equipped.

Figure 4, plate 22, shows the interior of a physical research laboratory in which gas-discharge studies are in progress. This laboratory is two modules or 12 ft. in width.

Figure 5, plate 23, shows the interior of a chemical research laboratory with an island bench. The services for the bench are extended to it from the outer wall by a trough under the floor that is integral with the building structure. This laboratory is three modules or 18 ft. in width.

Figure 6, plate 23, illustrates another feature of flexibility of the building design. This is the metal fabrication unit of our metallurgical laboratory. Because of the size of the machines, a very large space is required. This laboratory is eight modules or 48 ft. deep and 54 ft. wide. This width is obtained by placing the laboratory at one end of the rectangular block of a main stem and terminating the corridor eight modules from the end of the stem. This makes possible a laboratory extending across the full width of the structure. We have made similar use of the ends of the main stem rectangular areas in a number of situations to provide for laboratories of specially large area.

The flexibility of the laboratory space has made a large contribution to the operations of research and development. Rearrangement of laboratory rooms and facilities can be done expeditiously, with little loss of time, and at small expense. There is, therefore, no reluctance to rearrange space as the character of work changes. We can readily house in adjacent space the work on programmes where close relationships of the people involved are required. Development projects, by their nature, have small beginnings, and if successful expand as progress is made. The flexibility provided by our structure design makes possible the expansion of the working area of a development project as it matures.

The space for offices, conference rooms, drafting rooms, and clerical work is provided for in the wings that are at right angles to the main stems. This space does not require the large variety of special services that must be available to the laboratories. Economies were obtained by their segregation, which permits supplying those areas with only the limited services required by office space. The flexibility in room size has also been designed into these office and service wings. The same readily removable partitions are employed. The width of the module, however, has been increased from 6 to 9 ft., and the rooms on each side of the main corridor are 12 ft. in depth. The space can be subdivided into offices of different sizes, conference rooms, or large areas for drafting and clerical services.

Figure 7 is a diagram of a typical layout of laboratories in the main stem and of office, clerical, and drafting space in two service wings that are directly opposite each other across the width of the main stem.

DEVELOPMENT PROGRAMMES FOR THE MILITARY

The story of Bell Telephone Laboratories would not be complete without reference to our military development programmes. During the war we turned our attention almost completely to the creative technology of instrumentalities for warfare. New facilities for communications, radar, sonar, gun directors, etc. are typical of our areas of work.

At the close of hostilities, with military preparedness an important element of national policy, our Laboratories has continued to serve the nation in military creative technology.

The philosophy of work and organization of effort that have evolved in our technological services for an operating organization, the Bell System, makes us specially suited to serve the military, another operating organization. The pattern
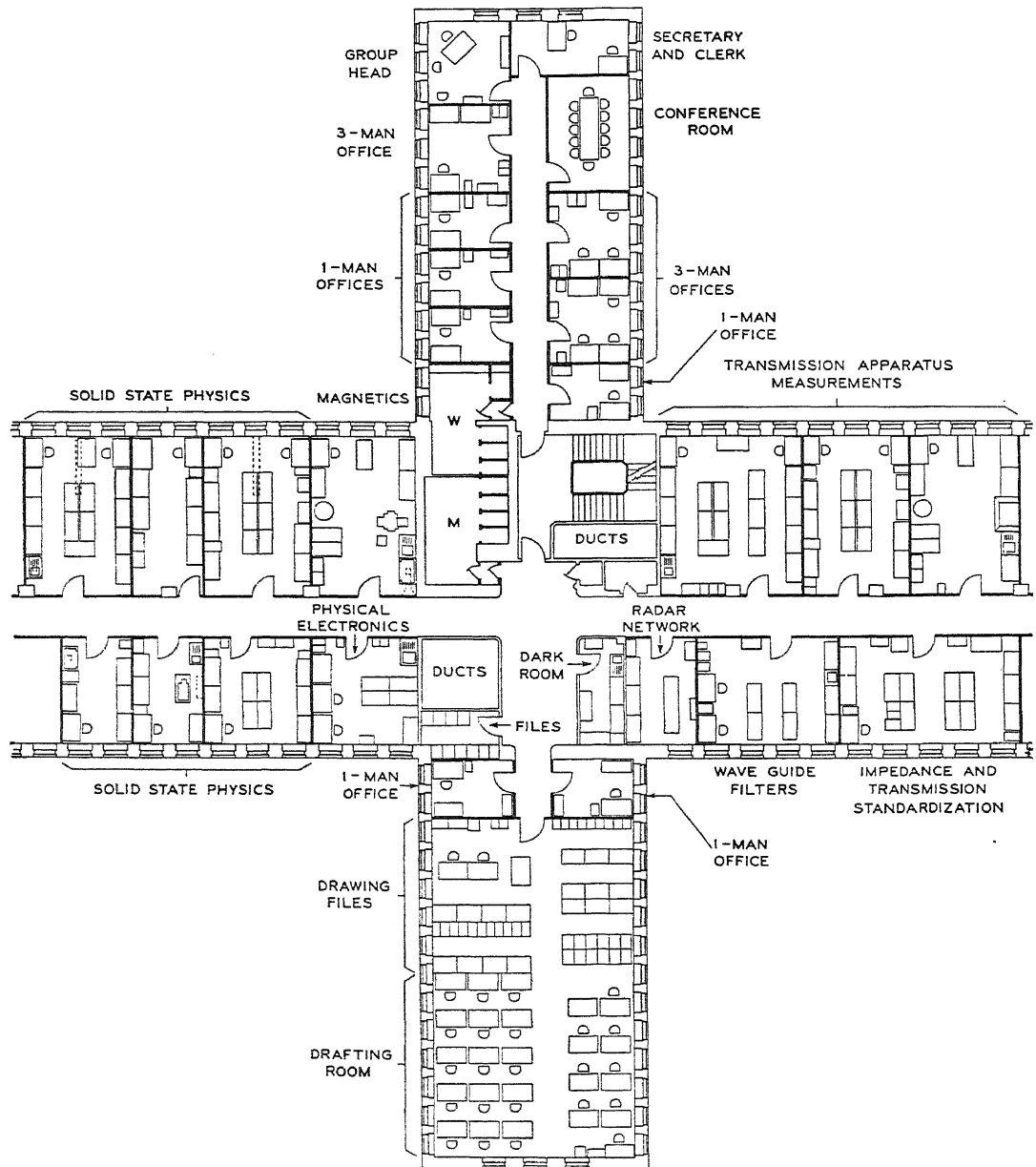
FIGURE 7. Diagram of a layout of a typical laboratory and office service area.

of our activities for the Bell System—a broad programme of research and fundamental development to obtain new knowledge of value to telephone service, the scientific determination of the best applications of the new knowledge in the

interest of service, and the development and design of new facilities best suited to service needs—is the pattern in which we are carrying out our work in creative technology for the military. The work on military systems proceeds in the same orderly manner through the successive steps from research to designs and specifications for manufacture and engineering practices for service which has proved so successful in the creation of new and improved systems for the services of telephony.

A discussion on the measurement of growth and form

UNDER THE LEADERSHIP OF S. ZUCKERMAN, F.R.S.

(*Discussion held* 16 *March* 1950—*Received* 17 *May* 1950)

CONTRIBUTIONS

THE PATTERN OF CHANGE IN SIZE AND SHAPE

By S. ZUCKERMAN, F.R.S., *Department of Anatomy, University of Birmingham*

In introducing a discussion on the subject of growth, it is unnecessary to add to the many definitions of the word which previous writers have provided. It is, however, useful to try to define the particular theme of the present discussion, for as Weiss (1949) has put it, the term 'growth' has become a cover for a variety of diverse and complex phenomena. 'It is not even', he writes, 'a scientific term with defined and constant meaning, but a popular label that varies with the accidental traditions, predilections, and purposes of the individual or school using it. It has come to connote all and any of these; reproduction, increase in dimensions, linear

increase, gain in weight, gain in organic mass, cell multiplication, mitosis, cell migration, protein synthesis, and perhaps more.'

The present discussion has been designed to illuminate an aspect of the subject which is aptly described by Richards & Kavanagh's (1945) definition of growth as 'a fundamental attribute of living organisms, manifested by a change in size of the individual, or in the number of organisms in a unit of environment', the change normally being an increase, with the possibility of negative growth or decrease in size under adverse conditions. Its main theme is the unfolding with time of a pattern of change in an organism—a pattern of change of size and of change of shape—and the methods by which the change can be defined. This was the aspect of the subject which dominated D'Arcy Thompson's interest. 'The study of form', he wrote (1916, 1942), 'may be descriptive merely, or it may become analytical. We begin by describing the shape of an object in the simple words of common speech: we end by defining it in the precise language of mathematics; and the one method tends to follow the other in strict scientific order and historical accuracy.'

By emphasizing theoretical issues relating to the description of growth processes, the present discussion is distinguished in plan from that of certain others which have been held in the course of the past five years. The Princeton Conference of 1946, a report of which has just appeared under the title *The chemistry and physiology of growth* (1949);* the 1947 Conference of the Society for Experimental Biology, *Growth in relation to differentiation and morphogenesis*;† and the 1948 Viking Fund New York Conference on the Study of Growth,‡ were largely devoted to studies of particular aspects of the subject, such as the synthesis of proteins, the genetics of pathological development, and the study of growth in man. Each was distinctive in its own way, and between them they cover ground which is essential to any full understanding of the subject. None of them, however, was concerned in any sustained way with the problems involved in the measurement of growth and form. By throwing this topic open for discussion, the present symposium extends the theme of some of the more theoretical contributions to the Festschrift published five years ago in honour of D'Arcy Thompson.§

Levels at which growth can be studied

Almost every discussion of the subject that has ever been held has brought out, at some stage or other, and in one way or another, the view that growth can be considered at the molecular level—in particular, at the level of protein synthesis; at the cellular level, which comprehends the process of cell division; at the organ or tissue level; and at the organism level. Each of these levels, if they can so be called, can be studied in terms of change in form. Each increasing level of complexity could in the abstract be regarded as based upon its predecessor. The interrelations and properties of its constituent molecules provide a statement of

---

* *The chemistry and physiology of growth* (1949), Princeton University Press.

† *Growth in relation to differentiation and morphogenesis* (1948), Cambridge University Press.

‡ 'The study of growth', in *Yearbook of physical anthropology* (1947), published by Viking Fund, New York.

§ *Essays on growth and form*, edited by W. E. Le Gros Clark and P. B. Medawar (1945), Oxford: Clarendon Press.

the form of a cell. The differentiation of a tissue or the growth of an organ is comprehended by the behaviour of its constituent cells—by 'their order of proliferation, their movements, transformations, affinities, secretions, etc., and the spatial arrangements and restraining conditions resulting from group occupancy of a common space' (Weiss 1949). And finally, the change in shape and size of an organism could theoretically be appreciated in terms of the growth of its constituent organs and tissues. While it is axiomatic that each so-called level of organization, and of change, should be explicable in terms of that of its predecessor, it is also a fact that our available knowledge does not permit us to describe the form of any one level by that of its predecessor. We can deal with each separately and in turn, but we cannot, starting with the molecule, define the form of the organism, nor, from a knowledge of the form of the organism, define the shape of the molecule. So far as the living thing as such is concerned, there are in fact no different levels of organization, for all the so-called levels are part and parcel of the same thing, on different aspects of which we can focus our attention at any given moment, as we so desire.

Laws of growth

It is useful to emphasize these points at the start because they bear *a priori* on an issue which will undoubtedly come up in the present discussion, namely, whether precise numerical formulations of temporal changes in size and shape can be derived from some fundamental biological law of growth; or whether numerical generalities are purely empirical measures which, in D'Arcy Thompson's words, in turn become steps in an analytical process. An answer to this question should emerge from the present discussion, which, starting at the documentary level of description, will proceed to a consideration of procedures of numerical formulation, and also consider whether these formulations lead us to a better understanding of problems of organization than would be possible without them.

The documentary phase of the problem

The accurate description of the changes which occur in the size and shape of an organism during its lifetime is in the first instance an anatomical matter. This is a field, however, where anatomy must inevitably use a numerical language or declare itself defeated. For 'the form of an object', as D'Arcy Thompson has put it, 'is defined when we know its magnitude, actual or relative, in various directions; and Growth involves the same concepts of magnitude and direction, related to the further concept or "dimension" of Time'. It does not matter whether we are discussing the growth of unicellular organisms, of structures in a multicellular organism, of a tree, or of a human being; if we wish to be precise about our impressions of change in size or shape, we have to use measurement and methods of numerical analysis.

The first step in the measurement of growth and form is therefore to obtain satisfactory data, in which the measurement of changes in form is related to time. Only when we have such data can we determine whether the character of growth differs between, say, a plant, a fish, and a human being. This documentary phase

of the problem is in itself of considerable interest, as it comprehends a mass of facts which are interesting by their novelty alone—facts about the longevity of animals, facts relating to the different phases of life of organisms, and so on. Is the way a plant grows fundamentally different from that of a mammal? Do trees die from what is understood by the term 'natural causes', or can they go on growing for what is called 'ever'? What is the extreme old age of fish, and do they go on growing endlessly? How 'prodigious' was the age of the Greenland halibut which D'Arcy Thompson records as having weighed 4 cwt., and as being nearly 2 ft. thick?

Growth curves

The proper expression of observations of this kind, and of the answers to the questions they pose, necessitates numerical formulation and analysis. Here we should distinguish at the start between the growth curve, which deals with temporal changes in any single dimension of the body, and curves which relate changes of shape to time. The dimensions which may be used to define curves of growth are the weight or any properly defined linear function of the organism. The characteristics of the primary curve of growth and its derivatives have been thoroughly explored by D'Arcy Thompson and, more recently, by Medawar (1945). The curve of growth can be expressed as an algebraic equation whose chief function, according to Medawar, is to facilitate analysis. 'Information about growth-rates and accelerations can be extracted from it by mathematical rule of thumb rather than by tedious and more or less inaccurate numerical or geometrical approximations.' The straightforward curve of growth can be readily translated into the curve of specific growth, in which not size itself is plotted against time, but the logarithm of size.

So far as is known, the general features of the growth curve are much the same in different organisms. 'Whether the animal be long-lived, as man or elephant, or short-lived, like horse or dog', writes D'Arcy Thompson, '...in all cases growth begins slowly; it attains a maximum velocity somewhat early in its course, and afterwards slows down...towards a point where growth ceases altogether. But in cold-blooded animals, as fish or tortoises, the slowing-down is greatly protracted, and the size of the creature would seem never to reach, but only to approach asymptotically to a maximum limit....' Conversely, 'among certain lower animals, growth ceases early but life goes on, and draws (apparently) to no predetermined end. So sea-anemones have been kept in captivity for 60 or even 80 years, have fed, flourished, and borne offspring all the while, but have shown no growth at all.'

Medawar (1945) has summarized the major conclusions that can be drawn from the curve of specific growth and its derivatives. 'Size', he writes, 'is a monotonic increasing function of age', and 'what results from biological growth is itself, typically, capable of growing....In a constant environment, growth proceeds with uniform specific velocity', whereas, 'under the actual conditions of development, the specific acceleration of growth is always negative'—by which he means that the specific growth rate is always falling, living tissue progressively losing the power to reproduce itself at the rate at which it was formed. This generalization, which was first stated by Minot, should, Medawar suggests, be known as Minot's

Law. And finally, 'the specific growth-rate declines more and more slowly as the organism increases in age'—a generalization which is summed up in Minot's epigram, 'organisms age fastest when they are young'.

The problems involved in the treatment of growth as change in size with time, and as expressed in the conclusions to which I have just referred, are far more widely appreciated and understood than are those which deal with the changing shape of an organism.

Differential growth

Here we have to make a clear distinction at the start between shape as a function of age, and changes in the relative proportions of different parts of the body. For time is eliminated from analyses of the latter kind. I should like to deal with them first.

The name of Julian Huxley is intimately associated with the question of relative growth. As he points out (1932), all but the simplest animals reach their adult shape and size by differential growth in different directions. His work has given prominence to the expression of change $y = bx^k$ or $y = bx^\alpha$, usually known as the allometry equation. This equation had previously been used to express a few specific organ-weight/body-weight relations in animals, but it is Huxley's achievement to have suggested that the simple allometry formula, which is usually given in the linear form, $\log y = \log b + k \log x$, can be widely applied to relate the growth of any part of an organism to the total growth of the organism as expressed, say, in its total length or weight. Specifically, the equation implies that the interrelation between the size of a part and of the whole is best expressed by a power function, and that the ratio between the specific or logarithmic growth rates of any two dimensions remains constant through the growth period.

The equation has been extensively used for the purpose for which it was advocated by Huxley, and its significance, both particular in the mathematical sense, and general in the biological sense, has been the subject of considerable discussion. In his recent paper with Reeve, Huxley (1945) has also indicated several of its shortcomings. For example, he has shown that a single straight line on the double logarithmic grid will not, in many cases, describe a change in proportions during the entire period of growth. He also emphasizes that the allometry equation ignores the factor of time (on the other hand, the equation indicates that at any given moment during the period of growth form is a function of absolute size), and he has pointed out the great difficulty of linking, by means of any particular formula, differential growth with growth in time. But, as he points out, 'with the data at present available, no appeal to time-growth formulae can be accepted as either a criticism or a justification of the allometry formula'.

This is an important point. Whatever its shortcomings, the fact is that the allometry equation seems to be a useful method of describing differential growth of the parts of an organism. What may emerge during the course of the present discussion is some idea of the reasonable limits of its usefulness.

Richards & Kavanagh (1945) point out that the ultimate criterion which governs the choice of a given method of curve fitting is whether its use produces any

significant addition to the fund of biological information, and they go on to say that the first purpose of most growth studies, and the principal one of many, is the demonstration of an orderly relation between the elements, structural or chemical, of the growing organism, and the description of the relation by a formula. In their view, the outstanding contribution of the allometry equation to the study of growth has been the demonstration that growth follows an orderly plan. They point out, however, that in one sense the exponent in the equation may be considered to be a correlation coefficient, in so far as it is the relation of two specific growth rates. In their view, however, simple linear correlation is of no particular value in describing differential growth. The question is, of what value is it? According to D'Arcy Thompson, the size of a part is in many cases directly proportional to the total size of an animal, and a simple linear function gives a better fit between the variables than the parabolic relationship implied by the allometry function. Is it possible at this stage of our knowledge theoretically to define the class of relation which is better expressed by the method of linear regressions than one which is suited to treatment by the allometry equation?

In so far as the treatment of data about the proportions of the body have a purely practical interest, it is possible to choose between the two methods according to the purpose for which one wishes to use the data; and in certain circumstances it will be more convenient to use methods of simple correlation, as Sholl, Shepherd & Vizoso (1949) have recently indicated. Suppose, for example, one wishes to obtain the proportions of the body for different age groups in a population, in order to provide basic data for the design of, say, chairs and tables. The information required will include such dimensions as sitting height, elbow height when sitting, sitting depth, and so on. It has been found in practice that it is a perfectly reasonable procedure to derive useful estimates of these subdimensions of the body from their known relations to such principal dimensions of the body as total height and weight. The precision of the estimates, in terms of their confidence limits, can be calculated with the help of well-established statistical procedures, and can also be refined by taking into consideration such other pieces of information as, for example, the economic status of the population concerned, and its geographical location. A question a general biologist would therefore like to have answered is when to use the allometry equation and when the method of straightforward linear correlation. How wide, too, is the boundary zone in which both methods overlap and are equally useful? Both clearly provide a means of summarizing masses of detail; both underline the fact that differential growth in an organism is orderly; and both can be used as a basis for prediction. Both also have the limitation that in any single instance they tell only about the scalar relationship of one part of the body to another.

Transformation in shape

Although it is true that both methods can be used to define any number of relations between subdimensions of the body (for example, tables have been constructed giving the regressions on total height and weight of dozens of the

dimensions of the human body), neither provides a comprehensive statement about changing shape. For this it is more convenient to resort to other techniques.

These have been developed by Medawar (1945), and are fundamentally based on D'Arcy Thompson's method of Cartesian transformations. This method, as his chapter on the subject, 'The comparison of related form', implies, was designed to show how one organ or animal viewed in two dimensions can be 'transformed' into a 'homeomorphic' form by inscribing its outline, in a system of Cartesian co-ordinates and then deforming it along simple and recognized lines—for example, by altering the direction of the axes, or the ratio of $X:Y$, or by substituting for X and Y some more complicated expression. By so doing, as D'Arcy Thompson demonstrated so dramatically, a new figure is obtained 'which represents the old figure under a more-or-less homogeneous *strain*'. In this way he showed how to transform the shape of the porcupine fish, *Diodon*, into the very different-looking sunfish, *Orthagoriscus*, and *Eohippus* into *Equus*—and, in fact, how to derive any number of artificial and imaginary intermediate types.

The method is essentially a static one, for it does no more than treat the shape of one animal, or part of an animal, as a geometric function of the shape of another. It is made dynamic by Medawar, who relates it to the growing organism, not, as he writes, 'with a view to reading a wonderful significance into any particular equations of transformations, but simply because the general analytic properties of the equations may be of first-rate importance'; and secondly, because the method can be put 'into motion' so as to express continuous relationships between infinite numbers of homeomorphic or related forms.

By this extension of D'Arcy Thompson's transformations, Medawar has been able to concentrate into a single equation such information as the fact that the legs get relatively longer, and the head relatively shorter, as we grow up, and that both changes are particular expressions of a single more general process of transformation. His method has shown clearly that change of shape 'keeps a certain definite trend, direction, or sense in time', and that the shape of the human being increases (or decreases) with age, the rate of change of shape falling off progressively in time.

Medawar's examples relate to subdivisions of the height of the body, for example, the distance from the soles of the feet to the fork, or to the navel, or to the chin, or to the top of the head. The method would seem equally applicable to values which comprehend both length and breadth. Furthermore, as Medawar has himself indicated, it provides a means of comparing the rates of change of shape of related organisms, for example, of apes and human beings.

A different technique which has been suggested (Richards & Kavanagh 1945) for studying change of shape, particularly in plants, is that of embedding in a growing tissue a small foreign body which does not grow itself, and which is sufficiently small not to interfere with normal growth, so as to follow its time-track within a series of co-ordinates. This method has been applied to the growth of leaves, but would at first sight appear inapplicable in the analysis of growth in animals—unless, of course, one were to take some such point as the navel in a human being as being equivalent to a foreign particle in a leaf.

Individual and mass curves

Before raising the question of the theoretical significance of growth curves, I should like to refer to one last practical point, to which an answer may be provided by the present discussion. Curves of growth may be derived from observations on a single growing individual, or a 'mass curve' may be constructed from the mean values for samples of individuals taken at different ages. Mass curves, it is well recognized, are corrupted by curves of distribution. This point is discussed in some detail by Richards & Kavanagh, who suggest that standard methods of correlation are more suitable than growth curves where it is necessary to rely upon averages. A question that arises, however, is whether an allometry equation derived from the means of randomly selected 'mass data' would provide a better expression of the relation between the growth of, say, the forelimb and the body than would the mean K value of a series of allometry equations derived separately for different individuals. No question of such choice enters, of course, into the consideration of the growth of internal organs. If we wish to know what is the relation between, say, the growth of the kidney and the weight of the body, we have to rely upon average figures for the weight of the kidney in relation to the weight of the body over the growing period. If it is more precise to avoid mass curves in describing the growth of the body, how is one to relate the picture of a 'mass curve' of the growth of an organ to that of the 'individual curve' of the growth of the body of which it is a part?

The theoretical background of growth equations

The issue to which I now wish to draw attention is the significance which can be read into equations that express growth changes in size and shape. Here there are two questions: First, can mathematical laws of growth be deduced from some fundamental law of behaviour of living matter? Secondly, can one read any biological meaning into empirical equations of growth? The first view appears to have few, if any, advocates to-day. As Medawar puts it, 'there is no biochemical entity to which the term *growth metabolism* may be applied; and no statement about it, therefore, which is likely to reward our confidence when it is placed (*qua* axiom) in a position of logical authority'. In so far as no single function characterizes biological growth in general, there can be no universal growth equation. If anything further were required to add to Medawar's denial of the view that growth equations can be derived from a knowledge of the behaviour of the living cell, it can be found in Weiss's more recent statement on 'Differential growth'.

However little *a priori* basis there is for the view that some universal growth equation can be deduced from the behaviour of living matter, it was inevitable that the lavish way in which the simple allometry formula appeared to demonstrate order in the differential growth of the body should stimulate the search for what has been called the theoretical basis of allometric growth. It was first believed that the logarithmic nature of the equation was a reflexion of the fact that living cells multiply by compound interest. This question has already been dealt with by Reeve & Huxley and by Richards & Kavanagh, and it is useful to refer to Reeve

& Huxley's conclusion of 1945, that 'no satisfactory theoretical basis has yet been found for simple allometry. The "axioms" of growth which were put forward to justify the general use of the formula are far from self-evident, and should perhaps be considered as no more than consequences of simple allometry where it has been found to occur.' In fact, as they add, we do not know how much of growth is multiplicative in the strictly mathematical sense.

The correct view would seem to be that there is no one fundamental biological law underlying growth equations, and that, to take an example, even if the transformation in shape of the human being can be summarized in one simple expression, this does not imply, as Medawar points out, that a single morphogenetic operator has been at work—however much it may be consistent with such a hypothesis. We can also agree with Weiss that a purely quantitative comparison of different growth processes is an abstraction. On the other hand, in doing so we should not conclude that it is a useless operation. The comparison of different equations of growth may draw attention to highly interesting differences in processes of growth. How fundamental, for example, is the difference between the way a plant unfolds, and the way a mammal does?

If, in spite of all that has already been written, equations of growth are not purely empirical procedures which, having summarized the facts of growth, allow one to derive general conclusions about the course of growth, we shall no doubt be hearing about it in the present discussion.

Differentiation and growth gradients

One stimulating concept which derived from Huxley's original treatment of the allometry equation was the conception of gradients of growth. The growth potential of an organ or region seems to be distributed in the form of a growth gradient with usually a single high point or growth centre, from which growth intensity grades downwards in both directions (or in one, if the growth centre be terminal). The existence of such gradients is fundamental to most processes of growth—not only to the shape of organisms but also of organs, and, as both Young (1945) and Weiss have so clearly shown in the case of the neurone, to cells as well. For example, the major zone of growth in the adrenal glands is probably the more peripheral part of the cortex, while the growing centre of a neurone is its cell body. Differential growth, as Weiss has argued, 'is simply a corollary of differentiation, and differences in growth rates result from a great variety of causes'.

The chemical environment of growth

Although it is not the main theme of the present symposium, it is improbable that the discussion will pursue its path without some reference to the molecular climate within which growth occurs, and the factors by which it can be disturbed. The rate of growth and the course of growth can be considerably affected by changes in temperature, in light, in nutrition (to refer to some of the constituents of the external environment within which the organism grows), or by relatively small changes in genetic constitution and in other internal factors which control bodily processes. Thus, growth is usually associated with protein synthesis, which can be

affected in many ways. Similarly, a slight alteration within a constellation of genes, or a change in a single gene, may manifest itself by multiple distortions of the normal course of growth—as, for example, in the various chondystrophies. Again, a slight alteration in the function of the thyroid gland may divert the normal process of development towards a cretinous goal. In other words, growth can be seriously affected if any one of a number of gears is changed. The unfolding of the normal pattern of growth is synonymous with a normal environment for growth.

Growth and growth rates in phylogeny

It would also be short-sighted not to take account, in the present discussion, of the use of growth equations in the study of phylogenetic problems. Simpson (1947), following Robb (1935), has emphasized this use, and exemplified it particularly in his discussion of the sequence from *Hyracotherium* to *Equus*. While the absolute rates of increase of muzzle length and of total skull length are different in this phyletic sequence, their relative growth tends to be constant, the relationship being expressed in the form of a single allometry equation. Simpson points out that practically the same equation applies not only to this phyletic sequence, but also in the comparison of existing breeds of horse and to the ontogenetic development of the horse of to-day.

While the allometry equation was not devised for the purpose of relating types dispersed through geological time, Simpson argues from the identity of the equation in phylogeny and ontogeny that the two varieties of muzzle length and skull length are genetically related, and that the striking changes that have occurred in the skull proportions of horses during the course of their evolution are simply a function of total size. Corresponding analyses of other characters have indicated changes in the relative proportions of different parts of the skeleton which it is assumed must have been due to genetic changes. And, in so far as the different individuals which come into the calculation are widely separated in time, Simpson's method makes it possible to estimate rates of evolution from this use of the allometry equation.

At first sight it looks as though, with the exception of the difference in time scale, the method is being used here in a way fundamentally similar to that for which it was devised, i.e. to analyse growth in an individual. Closer examination reveals many differences. First, in Robb's and Simpson's treatment, adults of different related types are being compared, not phases in the growth of a single individual. Secondly, and perhaps more important, there is an underlying assumption that the forms which are brought into the analysis are necessarily homeomorphic. That they may well be, but if they are, the conclusion is subsumed in the mathematical treatment, and does not emerge from it. A different method for estimating the rate of evolution from similar data that has recently been suggested by Haldane (1950) avoids the use of the allometry equation.

Growth and organization

It follows from the field of inquiry which I have briefly tried to outline that there is no single problem of growth, but as many as one wishes to create. Even the

problem of the measurement of growth and form appears multiple and complex, and the difficulties apply not only to the organism as a whole, but also to the smallest of its constituent parts which can be identified as such. The genetic, biochemical and nutritional problems which underline changes in size and form are innumerable. In short, the problem of growth could be taken to comprehend almost all the problems of biology.

Size and shape are clearly dependent upon growth, but growth, in the sense of the multiplication of protein and other molecules—or, to use Weiss's definition, 'the increase in that part of the molecular population of an organic system which is synthesized within that system'—comprehends much more, in so far as it also includes tissue replacement, regeneration and organization. At any given moment the form of an organism, organ or cell is to all intents and purposes static. On the other hand, cellular structure, as such, is labile and fluid—as has been impressed on us not only by older cytological studies, which pointed the conclusion clearly enough, but even more by recent metabolic studies in which radioactive substances have been used. It is, however, labile only within the limits set by the normal shape and size characteristic of the organism, organ, or cell.

Normal growth and regeneration thus display the unfolding of a pattern of organization, and the study of growth, starting with the fertilized ovum, provides one with the clearest picture attainable of the existence of such organization, and of the regulatory devices which, in the abstract at any rate, can be read into the developmental picture and the maintenance of normal function. The processes of growth and regeneration become the overt expression of that fundamental characteristic of living matter, dynamic self-regulation. In so far as the continuous and integrated co-ordination of bodily processes represents one of the central problems of biology, the analysis of the pattern of change in the size and shape of organisms will also continue to stand near the centre of the subject.

SECULAR CHANGES IN THE HEIGHTS OF BRITISH PEOPLE

By G. M. MORANT, D.Sc., *Air Ministry*

An inquiry has been completed recently involving comparisons of all the best records of heights of British adolescents and adults. They cover the period from about 1850 to the present day, and far more males than females are represented. The main object of the inquiry was to find out any changes there may have been in the age curve for height during the past hundred years. It is commonly supposed that the general situation regarding this question is known. The best records for British series of children are in good agreement in showing that height standards for particular ages were improving from decade to decade, if not from year to year. The latest generation was repeatedly found to have the highest mean heights for school years of age, and this secular trend was persistent. It has often been supposed that British people generally were becoming taller, so that the age curve for

a filial generation would be found to be above that for the parental generation throughout the whole age range from birth to the oldest ages. But this is an inference based on the records for children and it might be incorrect. An alternative possibility is that there was a secular change in the rate of growth but that more rapid growth was associated with the attainment of maturity at a younger age; the faster growing children when mature might be no taller or shorter on the average than their parents were. The question could be examined directly by comparing the records collected at different times relating to the last stage of growth and adult years of life.

The evidence has to be treated by making detailed comparisons of data for a considerable number of series and subseries of people. In this brief account of the inquiry it is only possible to indicate the more important general considerations involved and the principal results obtained. The records referred to are for a total of more than two million British men. The material is of a miscellaneous kind. It is due chiefly to various official and other bodies and to independent research workers whose activities were not co-ordinated. The need for a national anthropometric survey has often been advocated in this country, but one has never been established.

The series show diversity in various ways:

(a) The dates when the heights were recorded range from 1845 to 1948, and there are more series for the present century than for the last half of the previous century.

(b) The age range represented is different for different series. A large part of the material consists of records for recruits for the fighting services, or for men examined for that purpose whether they were accepted or not. The numbers are largest for ages 18 to 25, and this range usually covers the last stage of growth and includes the age when maturity in height is normally attained. Figure 1 shows a typical 'curve' given by mean heights for different ages. The data are for 216,000 men who were medically examined in 1917–18 in the west-midland region of England, with a view to selecting those fit for military service (Ministry of National Service 1920). The curve shows means increasing to a maximum about age 25 and then declining with advancing age. This form is shown by all the longest series. In making comparisons between different series an estimate is required for each of the ages at which the greatest mean height is found. The method used was to take the greatest mean and the two on either side of it and to apply a difference formula which gives the position of the maximum of a parabola passing through the three points (Yule & Kendall 1945, p. 487). Alternatively, the method may be applied to smoothed means, but the estimates obtained in the two ways are always close, as are the two shown in figure 1. The estimate of the age for maximum mean height given by a particular series will clearly not be of much value unless the means near the maximum of the age curve are based on fairly large numbers of individuals. As a rough guide it can be said that each of the means should be for at least 200 subjects. But even if the evidence is far more adequate than this there is still a danger that the estimate of the age for maximum height may be misleading. A particular series as a whole will be supposed to represent a particular population.

There is a danger that the way in which the subjects measured were selected was not the same for all age groups. Figure 1 provides an example of such lack of uniformity in sampling. The youths aged 18 had not been examined by recruiting medical boards before and they represent the total population of the region. In the case of all later ages many of the men had previously been examined and rejected for military service, and they represent the population of the region after withdrawal of considerable numbers of men who had joined the fighting forces.
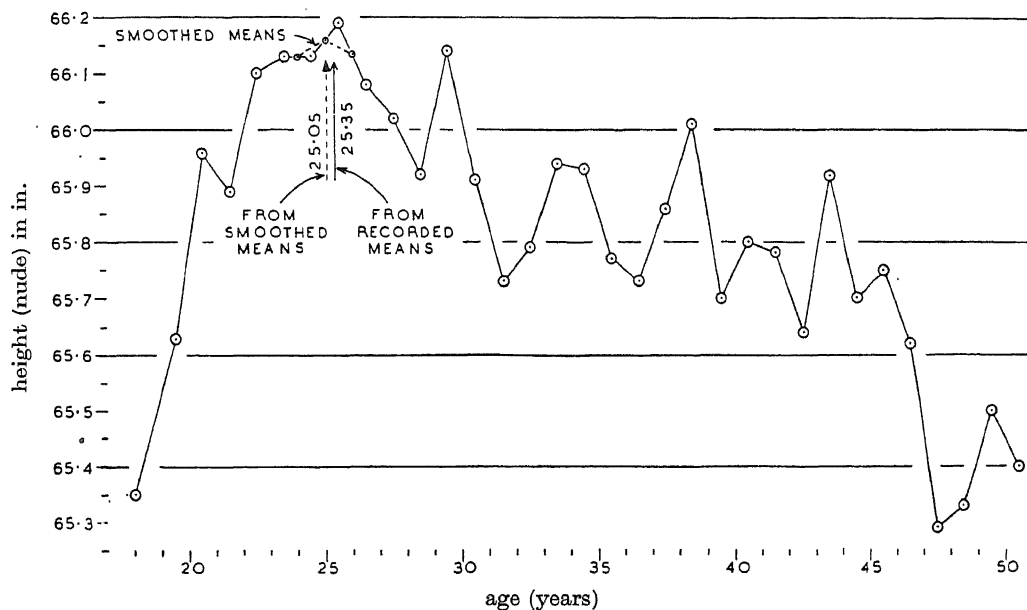
FIGURE 1. Mean heights at different ages of men examined by National Service Medical Boards 1917–18, showing ages for maximum mean height estimated from recorded and smoothed means. Vertical arrows indicate estimates of age for maximum mean height.

If for some unappreciated reason there was lack of uniformity in selection for means near the peak of the curve, then the estimate of the age for maximum mean height obtained would be misleading. In any case that criterion is only a make-shift. Information regarding the ages when growth in heights is completed could only be given ideally by repeated measurements of the same subjects taken at intervals over a suitable age range. There are no extensive records of such a kind. The best I have been able to obtain suggest that the distribution of ages for the attainment of skeletal maturity—judging from heights—extends from about the sixteenth to the twentieth birthday in the case of Royal Air Force apprentices to-day. They are a stringently selected group having uniform living conditions. Heights are available for several long British series of adult men without division into age groups. Some use can be made of these, but they are not good material.

(c) The civilian series are also diversified because they represent different sections of the total population. But clear-cut definition of a parent population described by a sample is the exception rather than the rule. One of the series

best described from a statistical point of view is that of visitors who volunteered as subjects in Galton's anthropometric laboratory at the Health Exhibition held in London in 1884 (Ruger & Stoessiger 1927; Elderton & Moul 1928). What part of Britain or section of the total population regarding social class, health or other conditions of life can such a sample be supposed to represent? The average heights of civilian series of British men cover a range of about 5 in., and they show an orderly sequence which seems to have been quite undisturbed by the fact that the records were collected at different periods. The ascending order is: criminals—patients in asylums—patients in hospitals—unemployed industrial workers—employed members of the labouring, artisan and industrial classes—estimates for the general population—members of the middle class—members of the professional class—members of the 'most favoured class', with pre-war students at Oxford and Cambridge at the highest level. Judging from the records for series of men of all ages pooled, there is no suggestion that the average heights for different sections of the population, or the average for the total community, were changing with time, but this is a matter which can be examined in a more precise way.

(d) A large part of the evidence consists of records for series of recruits and of men serving in the armed forces. The largest single collection is of recruits for the army for whom there were annual returns from 1860 to 1913, relating to a total of more than 1,700,000 men (*Army Medical Department Reports*). The period before 1905 is represented by thirty-three annual series of recruits, and for all except five of these the greatest mean height is for the oldest age group ('25 and over'). From 1905 onwards the greatest mean is for age 24 for 7 out of 9 years. The height limit applied in selecting recruits was changed on several occasions, but in any particular year that and other factors affecting selection applied equally to all men over 20 as far as is known. A more detailed examination of this material suggests that the age for maximum height for the recruits tended to become younger during the first decade of the present century, though up to 1913 it had not fallen below 24 years. The age for regular recruits for the Royal Air Force to-day is about 20 years.

(e) The records for series of men medically examined for the forces are of various kinds. There was stringent selection in some cases, as when height limits were applied. One of the recent series, however, gives the closest approach achieved yet to random sampling of the total British population of a particular age. The records are for more than 90,000 men aged 20 who were examined under the Military Training Act of 1939 (Martin 1949). Those accepted for service were called militiamen before the war started but the data are for all medical grades, including those of men judged to be unfit for service. The report gave for the first time an adequate picture of the regional distribution of stature in Britain. Using small regions and narrow subgroups of height—which is permissible because the numbers are large—shows a regular gradation with the highest mean for the south-east of England, falling values on passing westwards or northwards, lowest means for Wales, the north of England and the south of Scotland, and one of the highest values for the north of Scotland, though it is significantly less than that for the south-east of England. Other evidence regarding regional distinctions recorded in the present century conforms fairly closely with that of the militiamen. The

earliest surveys bearing on the question are those of Dr John Beddoe (1870), who collected the heights of men from various sources from 1861 to 1868. Some of his material has to be rejected because of biased selection. Doing so leaves totals for British men aged 23 to 50 of 5700 representing the general population, 1800 lunatics and 1250 criminals. The numbers are small for the purpose, but the three maps agree with one another in showing the greatest mean for Scotland and the lowest for the south of England. This is a reversal of the distribution for the militiamen. The comparison seems to provide clear evidence of a change in the geographical distribution of mean heights in Britain between 1860 and 1940. At the same time the range of regional differences was reduced. A plausible explanation of these changes is suggested by data given in the report on the militiamen. They show that the mean height of the men settled in the districts where they were born was 0·3 in. less than that of the migrants who had left their native places. There were population movements in the past hundred years from the north of Scotland (which formerly showed the greatest mean) and to the south-east of England (which now has the greatest mean).

In considering the question of a secular change in height by comparing the records for different series of people, the principal considerations which have to be taken into account are the dates of the series, the ages of the subjects, the regions and social classes or other sections of the population represented by the series and any other information regarding selection. All these factors have to be appreciated and allowed for in reducing the mass of evidence to a convenient form. The criteria derived from series which are of greatest interest in connexion with the problem of secular change are: (a) the age for the maximum of the growth curve, i.e. the age for maximum mean height, and (b) the value of the mean at that age.

Figure 2 shows the age for maximum mean height in the case of all the British series of men suitable to give estimates in the way described. A series, or a group of series, is represented by a histogram, and its height indicates the estimate of the age in question. The range of ages is from about the 19th to the 29th birthday. The series are divided into four groups according to the dates when the heights were recorded. They are also grouped so that those of similar kind—regarding the section of the population represented or special selection—are together. A detailed examination of the diagram, supplemented by the less satisfactory but extensive evidence of other series not shown in it, leads to two general conclusions regarding the age at which skeletal maturity is attained.

(1) At any particular time the age (or more properly the distribution of ages) at which maximum height is reached makes clear distinctions between different social classes, being earlier for more favoured, and later for less favoured, communities. When there is selection within a class the group which is physically fitter matures at a younger age—e.g. men accepted for the fighting forces compared with those rejected.

Differences between the ages in question for different series are clearly dependent on both the section of the population represented and on the period to which the records refer. By taking all the available evidence into account an estimate can be obtained for each period of the average age at which maximum height was

reached in the case of the general population. The second general conclusion is derived from such estimates. It is:

(2) For the general British population there was quite a marked change in the age for normal attainment of maximum height during the past hundred years. As far as can be seen it tended to become progressively younger throughout the period. The change was slow until 1939, and after that date it appears to have been definitely accelerated. The best estimates of the age for maximum mean height that can be given for the general population are 26 years about 1880 and 21·5 years about 1945. The secular change was more marked for the lower than for the higher classes of the population.
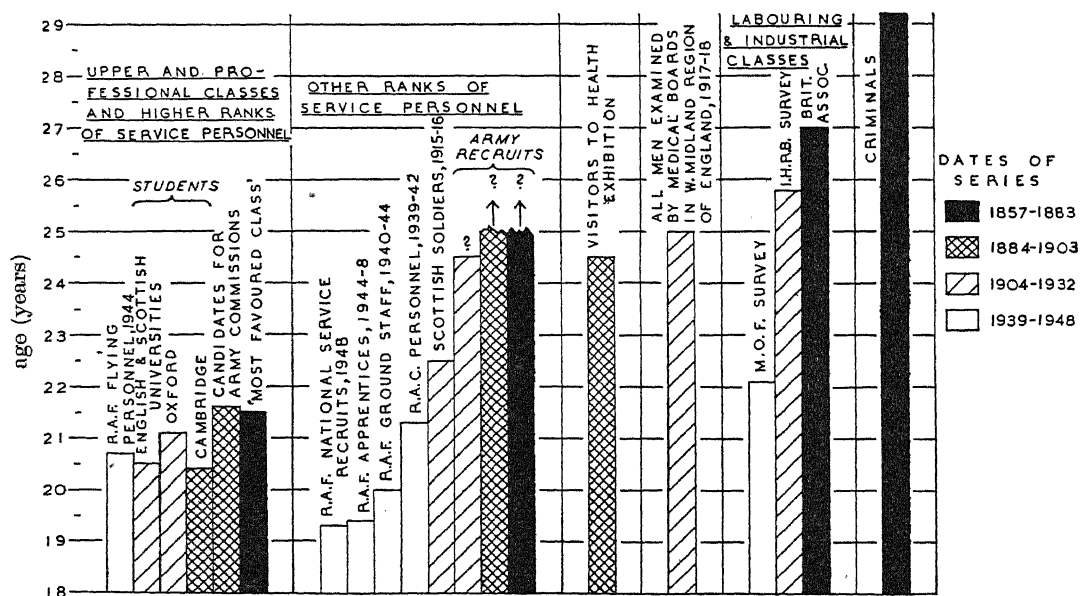
FIGURE 2. Estimated age of maximum mean height for various series of British men.

The other important criterion is the height level. In the case of series large enough to give good representation of the form of the age curve in the neighbourhood of its maximum, the most interesting index to use in comparing different series seems to be the maximum mean height, regardless of the age at which it occurs. In the case of a particular series a sufficiently good estimate of the maximum can be obtained by drawing a smooth curve through the points near the apex. Such estimates are shown in figure 3 for all the best series giving data for different ages. Indication of different periods is the same as in figure 2. For each period the values for all its series can be compared to give an estimate of the maximum mean height for the general population. For example, for the period 1939–48 there is a series of 27,000 male industrial workers measured for the Ministry of Food in 1943 (Kemsley 1950). The maximum mean height for this community (67·2 in.) is low. At the time most of the fitter young men of the total population were serving in the forces. The service series selected more stringently, though without applying height limits

—such as R.A.F. apprentices and flying personnel—have the highest means. Any value chosen as an estimate of the maximum mean height for the total population about 1943 would have to be between the value for the industrial workers and the lowest value for the service series, and 67·5 in. seems to be as good an estimate as any that could be given. Turning to the earliest period considered (1857–83), there are three civilian series representing the 'most favoured class', the labouring class and criminals. The best estimate for the general population may be taken rather above the value for the labouring class and close to, but possibly slightly above,
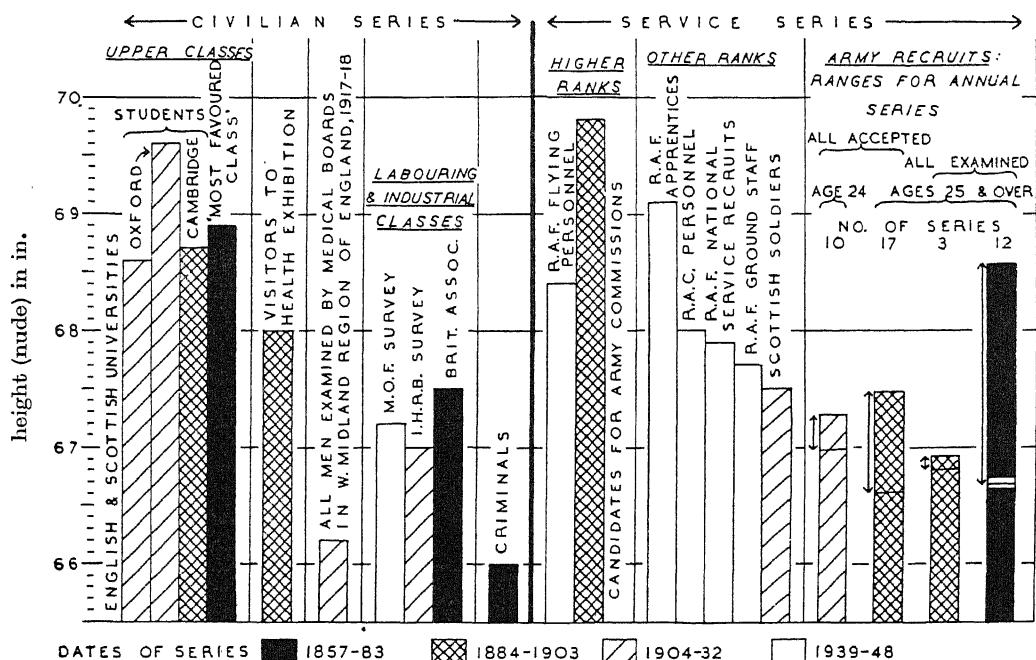
FIGURE 3. Estimated maximum mean heights for various series of civilian men and of men examined for or serving in the armed forces.

67·5 in. In obtaining estimates for the maximum height of the general population in different periods the evidence of many other series can be taken into account. These are not shown in figure 3, either because they give data for different ages but the numbers of men measured are not large enough to warrant the more precise treatment described, or because they do not distinguish age groups. Comparing all the material and when necessary making allowances as precisely as circumstances permit on account of age distribution, selective factors and regional representation, the third general conclusion suggested is:

(3) The best estimate that can be given for the maximum mean height of the general population of British men is 67·5 in., and this appears to have remained unchanged in the past hundred years. There are clear differences between the values for different classes of the community and these also appear to have remained unchanged.

Before summing up the situation as a whole it is necessary to refer to another conclusion derived from detailed examination of the evidence. This is:

(4) The normal situation is that the age curve for height for any community at a particular time shows a rise to a maximum somewhere between 18 and 30 years of age, followed immediately by a gradual decline which continues at a constant rate to the end of the age range. The rate of decline in adult years—of the order 1 in. for 22 years of age—can be supposed the same for the British series of all periods. The decline following the attainment of skeletal maturity can be attributed to a normal shrinkage of the length of the body with advancing age. It is probably
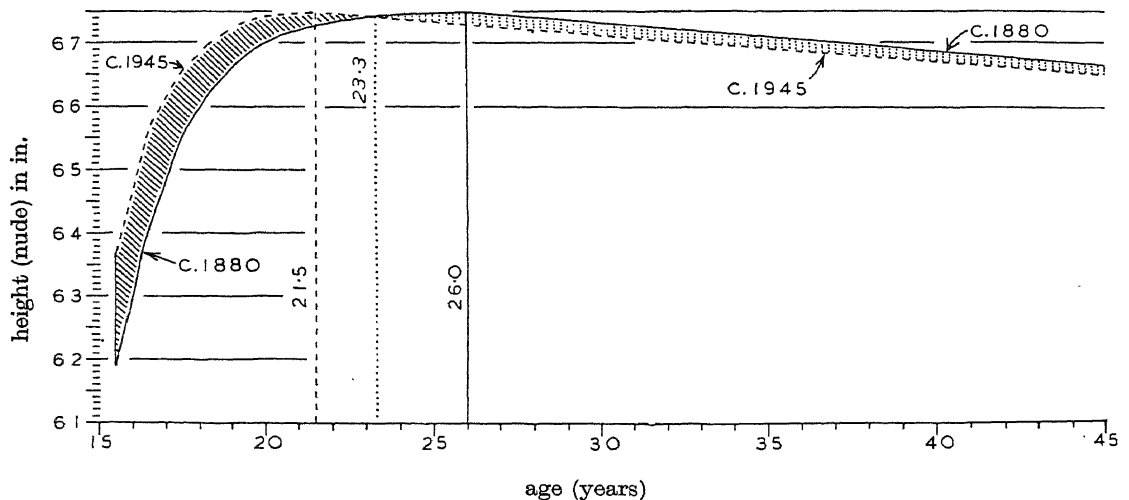
FIGURE 4. Idealized age curves of mean heights for the general British male population about 1880 and about 1945.

due principally to shrinkage of the intervertebral disks, but it may also be due partly to changes in joints and muscles affecting the erect posture. An alternative hypothesis explaining the decline as due to a secular change in the population— older men being shorter because they were born earlier—is not acceptable in view of all the evidence. The hypothesis that the decline may be due wholly or partly to a selective death-rate—taller men being supposed more prone to die at younger ages—is also unacceptable.

If the general conclusions given above are accepted then the situation regarding a secular change in the heights of the British males in recent times can be summed up in a single diagram. Figure 4 shows idealized age curves of mean heights for the general population about 1880 and about 1945. The two curves are shown as touching the same ceiling, which is supposed to be a height of 67·5 in. For the earlier period the maximum is at age 26, and for the later one it is at 21½ years. Accuracy in this feature to half a year of age is the best that can be expected. For particular ages the difference between the two curves is greatest at age 15½ when it is supposed to be 1·8 in. The gap is reduced to 1 in. at age 17, and it then declines slowly until the curves cross at age 23·3 years. After that the curve for the

earlier period is shown as being on top and the two run parallel courses, being 0·2 in. apart.

All the evidence is reasonably congruent in showing that there was an historical change of the kind illustrated, though details of the picture may not be quite correct. The secular transformation of the earlier curve until it eventually became the later one was evidently due to a speeding up of the rate of growth. As far as can be seen the movement throughout the period was always in the same direction —the age for normal attainment of skeletal maturity becoming progressively younger—and it appears to have been most rapid during the last stage, since about 1939.

The conclusions are that British people to-day are taller, on the average, at all adolescent ages than their ancestors were two generations back, but they become fully grown at a younger age and their maximum height is the same as that of their ancestors who matured more slowly. As the prime is reached to-day at a younger age than formerly the decline in adult life also starts earlier, and this implies that for ages after the early twenties people to-day are slightly less tall, on the average, than their recent ancestors were. The British records show that the maximum mean height for a population may be stable while the age at which it is normally reached is unstable. The statistical evidence suggests the heights which people have when they are fully grown are determined almost entirely, if not entirely, by heredity, but that growth rates within wide limits are determined by conditions of life. But it is for geneticists to provide the ultimate solution of a genetical problem. Conclusions derived from the indirect evidence may serve their purpose to some extent by emphasizing the need to make a clear distinction between the situation for immature ages and growth rates on the one hand, and that for the adult stage of life on the other.

Granting the conclusion that the period of growth in height became progressively shorter in the past hundred years, there are several implications of interest. Was the attainment of skeletal maturity at a younger age associated with parallel changes for other characters and faculties or not? What are the merits or demerits of speedier growth? There are other issues of a more technical nature, such as the following.

A number of body measurements commonly recorded for living people—such as stature, sitting height and the lengths of the limbs and their segments—are indirect measurements of the size of the skeleton. In an individual they become set, as it were, when he reaches his prime, though some, like stature, may show slight decline with age in adult years. Measurements of the class referred to are essentially different in nature from weight and girths of the body, which may change appreciably in either direction at any stage of life. It is known from experience that correlation coefficients between pairs of skeletal lengths are practically constant for different adult age groups of the same series, and a value of the same order is found if data for all adult ages are taken together. For adolescent ages variation in the measurements of length is found to be appreciably greater, and correlation coefficients are also appreciably higher. The distinction can be supposed due to the fact that at a particular immature age the individuals

in a sample represent a range of stages of development, but for all adult years the stage of skeletal development is practically the same. If there is secular change in the growth rate for skeletal characters, then it is to be expected that there will also be secular change in the whole system of interrelationships between the characters charted on an age basis. Owing to this fact it is more difficult to establish suitable age standards of body size. To do that with assurance the evidence is needed of recent surveys and the trend of secular change must also be taken into account.

MEASURING GROWTH IN FARM ANIMALS

By J. HAMMOND, F.R.S., *School of Agriculture, University of Cambridge*

There are various ways of measuring growth, as actual weight, percentage increment, or weight per day gained. Which method is used will depend on the purpose for which it is required. For agricultural purposes the third method is generally adopted, as it forms a basis for daily food requirements and, conversely, also enables the rate of growth to be controlled exactly by regulation of the food intake.

The live-weight growth curve of the animal as a whole is made up of the sum of the different parts and tissues of the body, each of these growing at different rates.
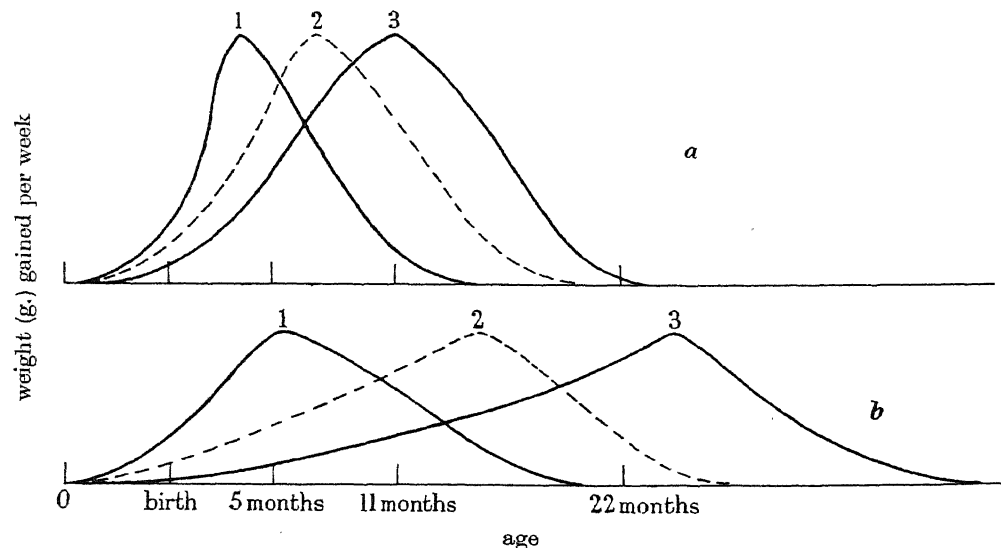
FIGURE 5. The rate of increase in weight, showing the way in which the changes in shape and proportions of the body are affected by the level of nutrition. *a*, on high plane of nutrition; *b*, on low plane of nutrition.

curves 1 { cranium and shanks
bone
caul and gut fat

curves 2 { neck
muscle
subcutaneous fat

curves 3 { loin
fat
marbling fat.

This gives rise to changes in the form and composition of the body as the animal grows up. Thus, as shown diagrammatically in figure 5, the maximum growth rate for bone is attained before that of muscle, while fat is latest in development. Similarly for different parts of the body, the head matures early, the neck later
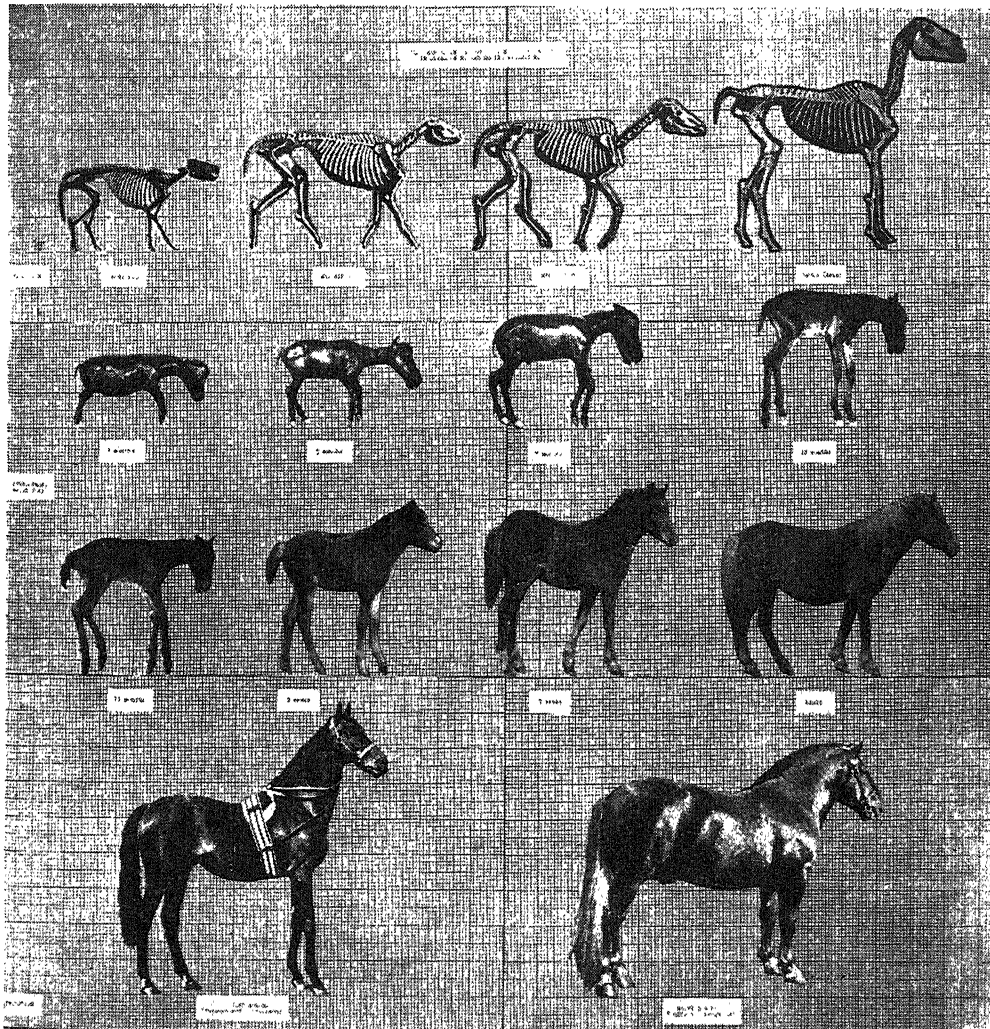
FIGURE 6

and the loin last. In late-maturing breeds and under low-plane feeding the curves for growth of the different tissues and parts of the body are extended in time, whereas in early-maturing breeds and under high planes of nutrition they are compressed (figure 5).

The main requirement in farm animals is to obtain an overall measure of the changes taking place in the form of the animal as it grows up, and for this purpose a photographic system has been adopted (see figure 6). In doing this it is important

to adopt as a standard for the basis of comparison an early developing part of the body, as then an orderly change in the proportions of the other parts will become evident. Moreover, such early developing parts are least affected by the plane of nutrition on which the animal is reared or by other environmental conditions. For example, when the overall changes in form during evolution and embryonic life are being compared, an early-developing part such as the brain weight, cranium or eye to ear length is particularly useful, for then all the changes in form up to the time of birth in the horse (figure 6) proceed in one main direction (relative length of leg for speed), while after birth a different phase of development begins (lengthening and deepening of the body for power), the Thoroughbred and the draught-horse respectively being extensions of these two phases of growth changes in form (Hammond 1940).

Another example of how, by selecting as a basis for comparison an early-developing part of the body such as head size and animals on a low plane of nutrition, it is possible to obtain an overall picture of the evolution of form in the domestic pig is shown in figures 7 and 8. Using head size as a basis for comparison, figure 7 shows how, in comparison with the wild boar, first the shoulders and legs and lastly the loin increase in proportions as the improved domestic pig grows up (Hammond 1947).

When an improved breed of pig is reared on a low as compared with a high plane of nutrition, and the weights of the different parts of the body at 16 weeks old in the high-plane animals are compared with the weights of the respective parts in the low-plane animals (figure 8, McMeekan 1940–1), it will be seen that the parts of the body are affected by nutrition in the same relative order as they have developed in evolution; the head is least affected, then the shoulders and legs, while the loin shows the greatest effects.

When postnatal changes only in body proportions are being measured, the shoulder height forms a very convenient standard, for its maximum proportions are attained at about the time of birth in farm animals, and consequently the changes with age, breed and sex are all shown in the same direction (figure 9, Hammond 1932); those changes of form which occur before birth (see embryo) will, however, appear in the opposite direction and as has been pointed out, for these a very early developing part such as the brain or cranium is preferable. Similarly, when postnatal changes in the composition of the body are considered, bones such as the cannons (metatarsals and metacarpals), of which there are four, giving a safe average, give a good standard for comparison because they reach their maximum proportions at about the time of birth; their weight is less affected by nutrition than other later developing bones of the skeleton, and very much less than the muscles or fat. If this is done then the true biological significance of differences between the growth rates of different parts and tissues in early- and late-maturing breeds becomes evident (see Hammond 1932).

Since the maximum growth rates of the different parts and tissues of the body occur at different times during the animal's life (see figure 5), it should be possible to alter the proportions and composition of the body by controlling the plane of nutrition at different phases of the animal's life. This has been done by McMeekan

(1940-1) for the pig by making the animals grow along predetermined growth curves (figure 10) by controlling the plane of nutrition. While his high-high and low-low pigs (figure 11, *H.H.* and *L.L.*) at the same body weight (200 lb.) show
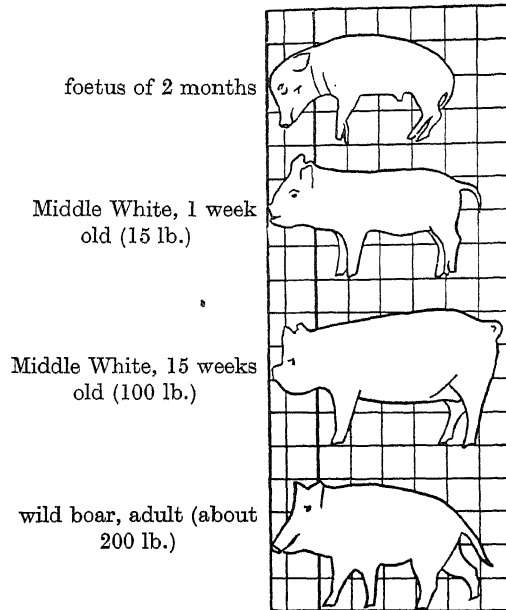
FIGURE 7. Changes in the proportions of the pig brought about by selection. Each animal is reduced to the same head size. As an improved breed such as the Middle White grows up, the proportion of loin to head and neck increases greatly: but an unimproved type such as the wild boar grows up without much change in body proportions. Compare with figure 8.
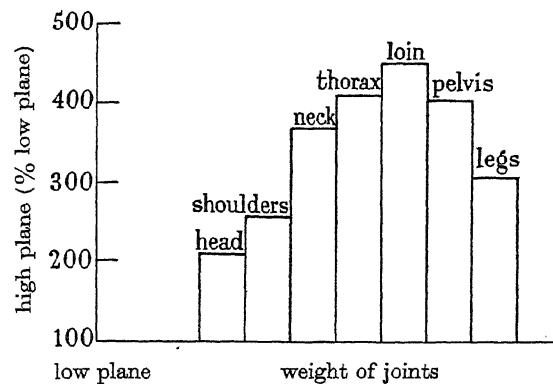
FIGURE 8. Plane of nutrition—body proportions at 16 weeks.

much the same changes as do pigs of the same age on high and low planes of feeding (figure 8) they are of very different ages. His high-low and low-high pigs, however, are of the same body weight (200 lb.) and same age (figure 11, *H.L.* and *L.H.*), and yet they show considerable difference in body form (figure 11) and
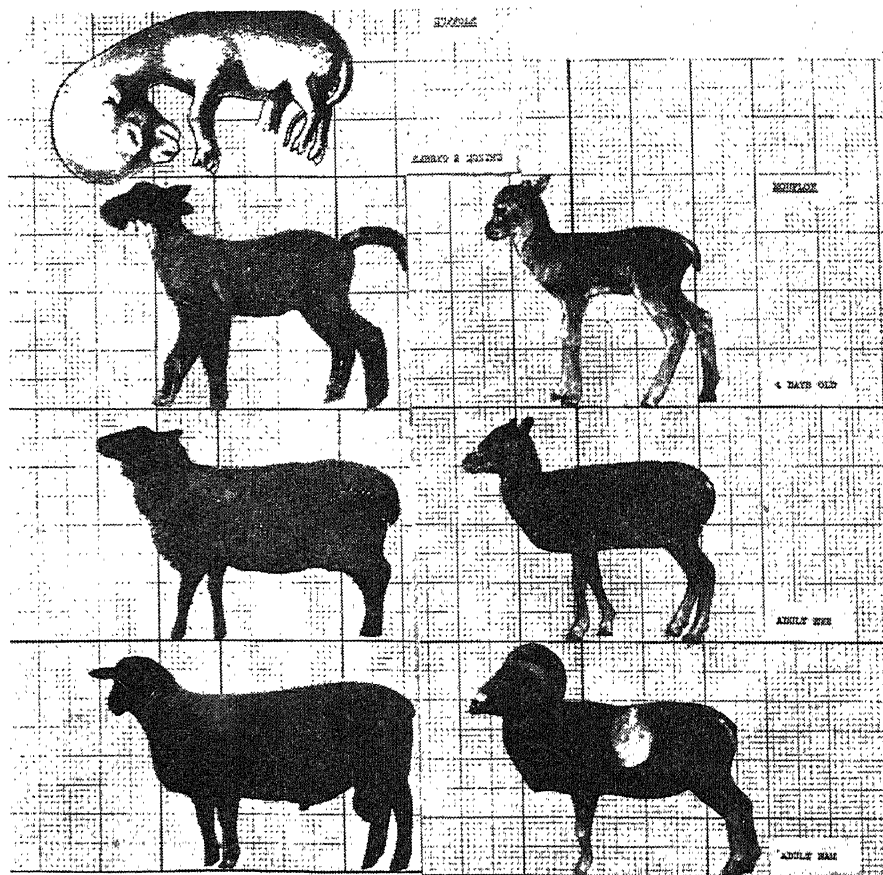
FIGURE 9. Change in body proportions (to same shoulder height) with age and sex in improved Suffolk and unimproved Mouflon sheep.
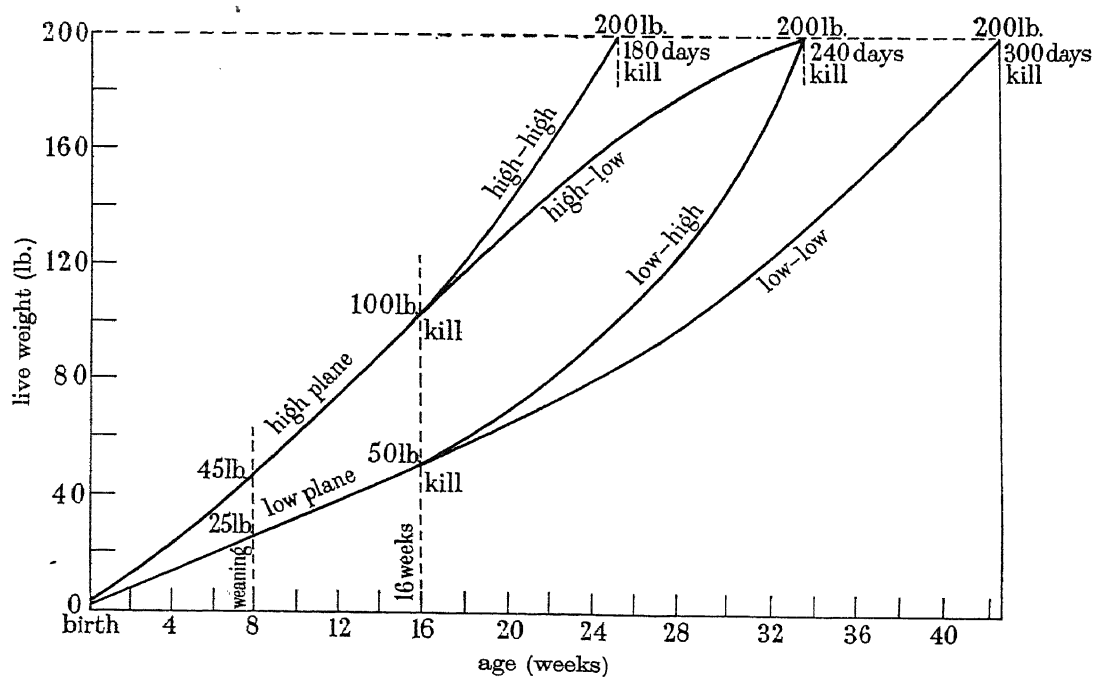
FIGURE 10. Plan of experiment. Live-weight growth curves to be secured by quantitative control of the plane of nutrition. (Differences in shape of growth curves between all treatments to be accentuated if possible.)

proportions of tissues (figure 12), because the shapes of their live-weight growth curves are different. Pigs fed on a high plane of nutrition when they are young, at

high-high
168 days old

high-low
196 days old

low-high
196 days old

low-low
315 days old
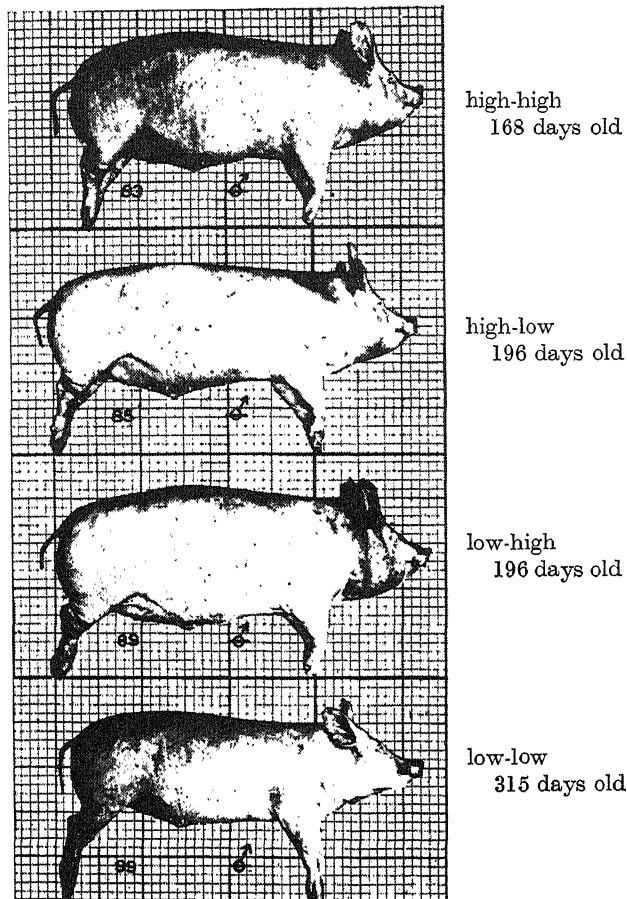
FIGURE 11. Pigs at 200 lb. live-weight. All to similar shoulder height.
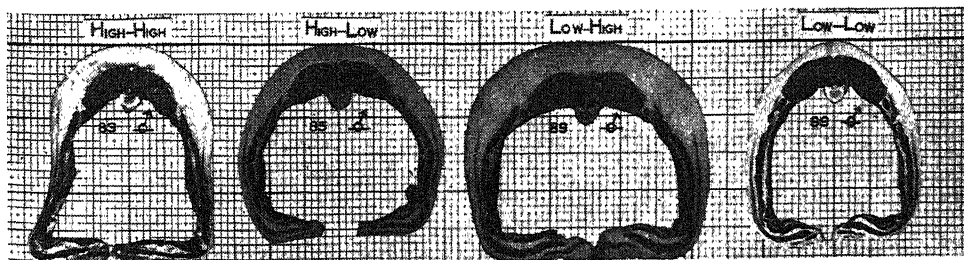
FIGURE 12. Section through pigs shown in figure 11.

a time when bone and muscle are growing quickly, develop a large frame, and when dropped to a low plane of feeding later, fail to grow fat (*H.L.*, figures 11 and 12); whereas in pigs which are fed on a low plane when they are young, the skeletal

frame is stunted, and when later put on a high plane, deposit large quantities of fat on a shortened body (*L.H.*, figures 11 and 12).

Such studies of the control of growth and form by nutrition lead to the formulation of a theory (Hammond 1943) of the partition of nutrients to the different tissues of the body according to the order of their development in the individual
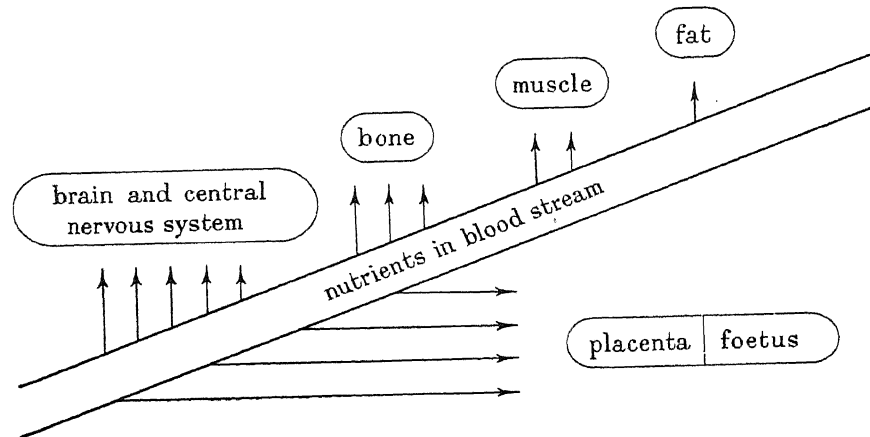
FIGURE 13. Priority of partition of nutrients according to metabolic rate.
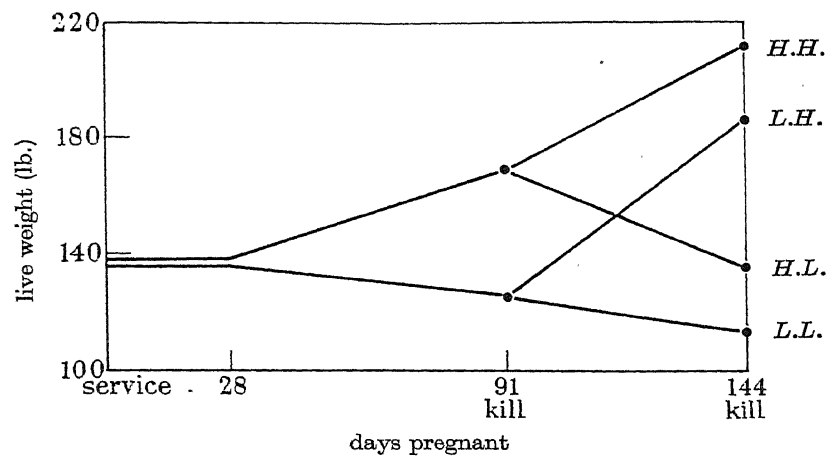
FIGURE 14. Live-weight growth curves of ewes during pregnancy.

or to their metabolic activity (figure 13). Thus the brain and bones of a young animal will continue to grow even though it is kept at a constant live weight over a long period of time, the necessary nutrients being obtained by withdrawals first from fat and then from muscle.

Within such a system during pregnancy, the placenta in the early stages has high priority, but as it matures toward the end of pregnancy this priority is lost, and so its nutritional uptake is then very dependent on the plane of nutrition of

the mother (Wallace 1948). By controlling the body weights of ewes during pregnancy by the control of food intake (figure 14), he showed that the growth of the lamb during the later stages of pregnancy, but not during the earlier stages, is
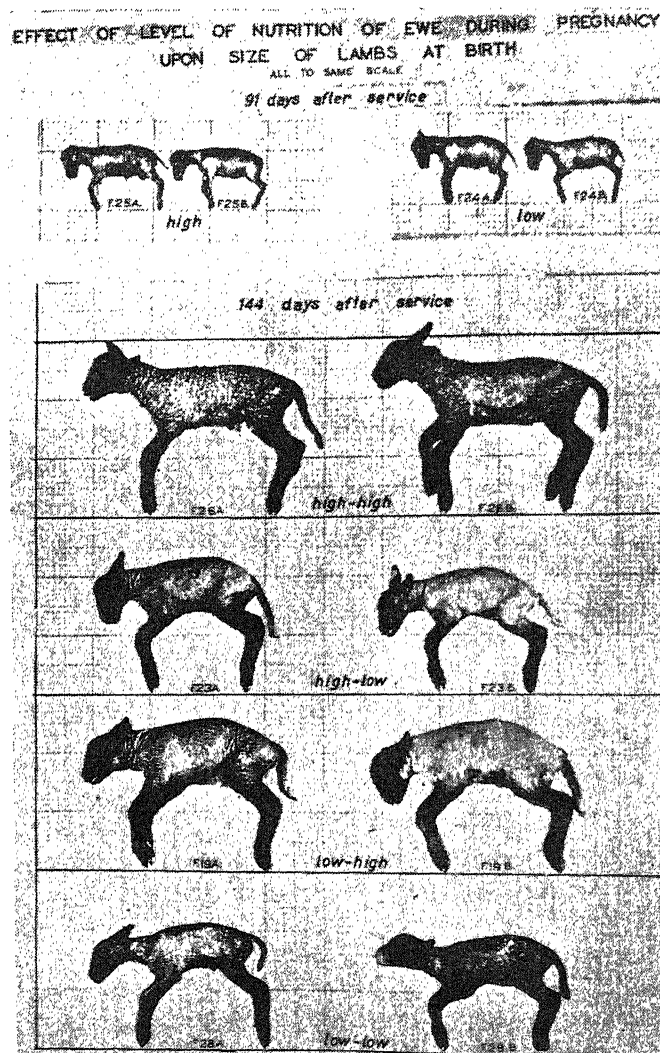
FIGURE 15. Twin lambs produced by ewes on the planes of nutrition shown in figure 4.

controlled by incoming nutrients (figure 15). Although all these lambs are of the same conceptional age (144 days from fertilization) they are of very different sizes and of anatomical and possibly physiological ages. As will be seen from figure 16, the larger ones have been advanced and the smaller ones retarded in the proportions of scapula to total skeletal weight, although the pattern of development has not been changed. In order to measure the relative effects of age as such compared

with actual body weight on the rate of ossification of the bones and other physiological age changes, the three-dimensional diagram as used by Appleton (1929) for newborn rabbits (figure 17) has many advantages. It shows that such growth in development depends not only on size but also on age. Hence the possibility of the control of body form by the control of nutrition at different phases of the animal's life.
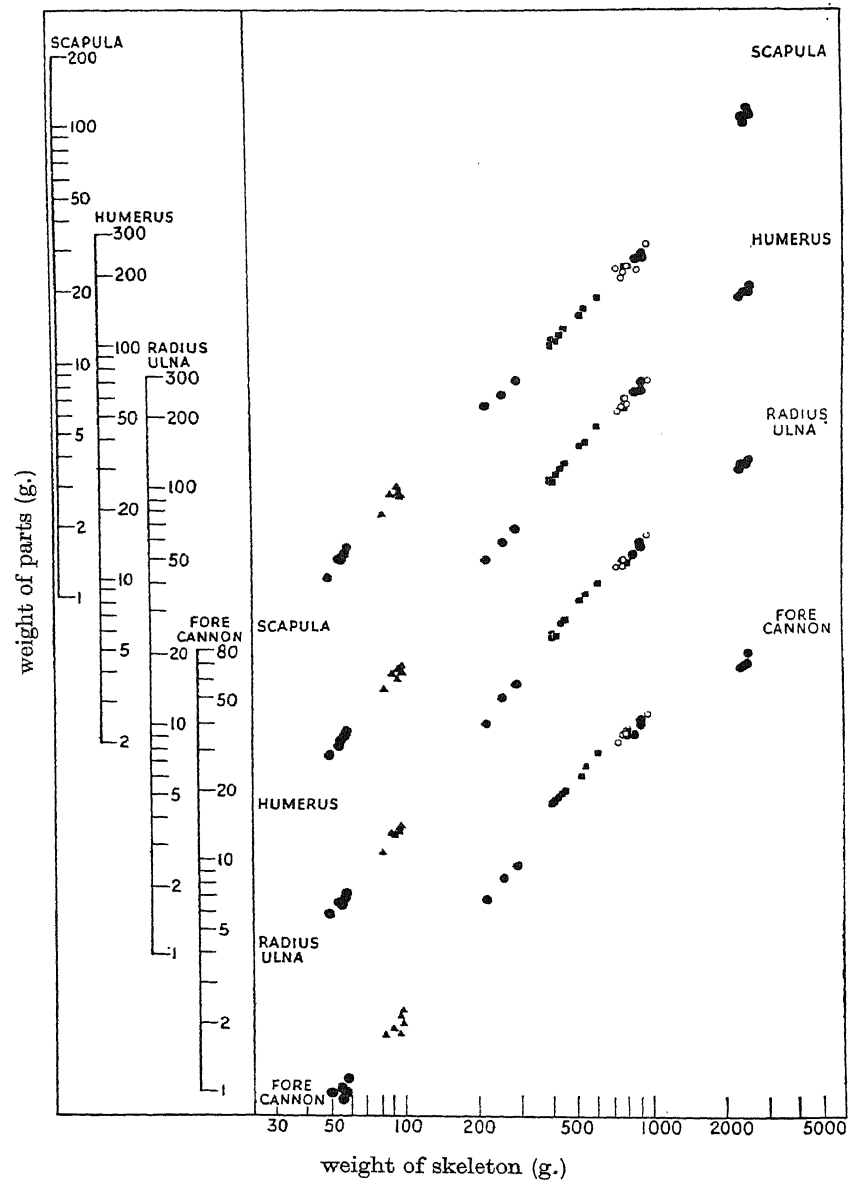
FIGURE 16.
● B.L.-Ch. × Suffolk.   Age series.
▲ Suffolk.   91 days.
■ Suffolk.   Low plane.   144 days.
○ Suffolk.   High plane.   144 days.
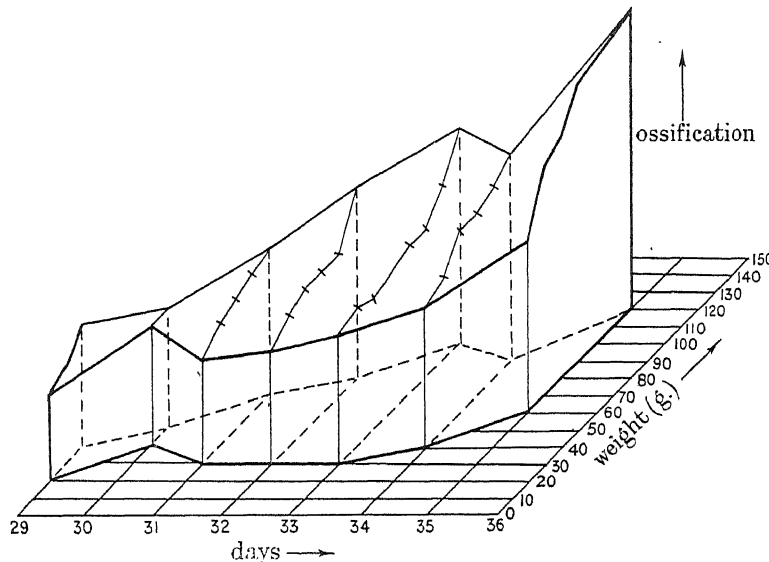
FIGURE 17. The relation between the stage of ossification in the limits of the new-born rabbit and its weight and age are shown as a three dimensional chart. The progress in the ossific condition as weight and age respectively increase is shown by the form of the upper surface of the solid figure.

GROWTH AND FORM IN PLANTS

By F. G. GREGORY, F.R.S.

Department of Botany, Imperial College of Science and Technology, London

There are fundamental differences in the growth processes of higher animals and plants. In the animal the phase of embryonic development during which the form of the animal is determined is of relatively short duration, whereas in the plant embryonic regions persist, and a change in form is therefore prolonged throughout the whole life history. A secondary marked change in form occurs in the higher animal at or around puberty, but in the plant puberty is not epochal, if, indeed, it can be said to exist at all. Again, in the plant, the form is very open to modification by external factors, such as intensity, quality, and duration of light, nutritive factors and the like. So far as the effects of external factors are concerned one has to distinguish between the immediate effects and after-effects which may predetermine the course of development to a remarkable degree. The phenomenon of vernalization is a striking example. By merely lowering the temperature during germination of winter cereals to the neighbourhood of the freezing-point, the process of initiation of flowers and development of the ear may be advanced by as much as six months, and the plant altered in its behaviour from that typical of a biennial to that of an annual plant. Again, the phenomenon of etiolation consequent upon keeping the plant in darkness completely alters the form of many

flowering plants. It does this by enhancing the elongation of the stem while leaf expansion is arrested.

The external form of the higher plant is in the main determined by four characteristics:

(1) The relative development of stem and leaf.

(2) The tendency for secondary axes or branches to be produced.

(3) The absolute size and shape of the leaves, and their arrangement on the stem.

(4) Duration of the growth of the primary and secondary stems.

I shall say a few words at first on these characteristics.

(1) *Relative development of the leaf and stem*

The preponderance of development of the stem is seen in the tree habit, which is associated with a persistent embryonic region in the stem, the cambium, which is responsible for secondary thickening of the trunk. In general, it is accompanied by the massive development of wood laid down as annual rings. This habit is not universal, as is seen, for example, in palms or tree ferns in which no secondary thickening occurs as the cambium is absent. The same habit is seen in shrubs which are merely diminutive trees. In herbaceous plants great variation is seen in this respect, and we find at the other extreme rosette plants in which the stem is virtually absent. In complete darkness even such plants show a marked elongation of the stem accompanied by a reduction of the leaves to mere scales.

(2) *Secondary axes*

Even without leaves trees are recognizable by their branching habit—everyone can distinguish between the graceful form of the birch and the stocky appearance of the oak. Not only the number of branches vary but also their relative development as compared with the main stem, and the angle at which they grow out.

(3) *The size and shape of leaves and their arrangement on the stem*

At this point I merely wish to draw attention to this character and shall return to it later. Apart from floral structure the leaf shape is probably the most definite form characterizing the species, though even this is modified by external and nutritive factors. Thus some aquatic plants, such as the arrowhead, produce two entirely different types of leaves. The submerged leaves are long and linear, the aerial leaves are arrow-shaped. This change in leaf form is in many cases associated with the spectral composition of the light in which the leaves develop. Again, in some succulent plants, the mere variation of day length modifies the leaf from a typical fleshy form without stalks to thin stalked leaves. Even in annual plants successive leaves differ in shape, and these differences in form are accompanied by variation in chemical composition and are therefore nutritive in origin.

(4) *Duration and growth of the primary and secondary axes*

In some perennial plants the stem growth persists indefinitely either with branches, or as in palms, without branch formation. In others the main stem persists for only one year, and is replaced in the following years by new branches

arising from underground stems. In the true annuals no such replacement occurs and the plant dies after a variable period of growth. This is associated in annual plants with flower and fruit production. The embryonic growing points are converted to inflorescences which terminate further development. Even in such cases by altering the length of day a plant can often be induced after fruiting to produce a new shoot from a dormant bud, and so repeat the cycle of development. Most plants have such dormant buds which can be induced to grow by cutting back the plant. The mechanism is hormonal; so long as the terminal region is active the buds laid down in the axil of each leaf are inhibited from growing. The degree of this so-called 'apical dominance' must be a factor determining the extent of branching of all plants.

The study of the form of plants

Bearing these considerations in mind we may approach the problem of the quantitative study of plant form. We may I think distinguish two aims in such a study: (1) a study of a circumscribed problem such as the changing form during development of a single organ such as the leaf, and (2) the study of the interrelated growth processes of the various organs of the whole plant. The study of the root presents peculiar difficulties when plants are grown in soil or other solid medium, though the physiologists have developed methods of growing plants in culture solutions, a practice which has found application in hydroponics.

D'Arcy Thompson, in his great pioneer work, has dealt with the question of the ultimate form of the leaf, and has shown how many of the purely descriptive names of leaf shapes used by a systematic botanist can be translated into mathematical terms. This is done by considering growth rates distributed radially about a node or 'point of arrest' at which no growth occurs. It must be admitted that a purely formal study of this kind in fact throws no light upon the actual mode of growth of the leaf. The shape of the leaf is quite independent of cell shape or size and is achieved by local variations in cell division, and appears to be related to distribution or synthesis of hormones. The hormone auxin locally produced during cell multiplication appears, however, to be a regulator of stem rather than leaf growth; no indisputable evidence for a leaf-growth hormone has yet been put forward. Such mathematical expressions may be of great interest in comparative morphology, and probably also in enabling precise statements of the effective external factors on leaf shape to be made. Naturally only relatively simple leaf forms can be studied in this way.

So as far as the second aim, the interrelation of the growth processes in the different organs, is concerned, considerable progress has been made. The question here is to select characters for measurement which may form the basis of an analysis of the developmental progress of the plant. Obviously it is advantageous to select characteristics which reflect the factors concerned in determining the changing form of the plant already mentioned. This is the method of growth analysis which has been used in relating the factors, climatic and nutritional, controlling growth to yield in plants. So far only annual or biennial plants have been studied in any great detail and mainly in association with agricultural

problems. A pioneer work of this kind was carried out by Crowther in connexion with the growth of cotton under irrigation in the Sudan and yielded results of importance both theoretical and practical.

The primary data collected are: (1) the height of the plants; (2) the increasing leaf surface; (3) the number and dimensions of the nodes on the stem; (4) the number and development of branches; (5) the number of flower buds and later of fruits, and (6) the total weight of the plant. The measurements can be made on the plant during growth and the total weight by periodical sampling. In this way growth curves of the whole plant and its constituent organs can be obtained and their interrelations studied. Since in the plant the material gained during growth accumulates, there being no excretory process of any magnitude, the growth curve is of primary importance, but there is no time to deal with this matter now.

Knowing the leaf area at any time and the weight increase during a given period, the carbon assimilation of the plant can be ascertained and the relative contribution to dry-weight accumulation of changes in the rate of expansion of the leaf surface and of changes in carbon assimilation can be assessed. By chemical analysis of the plant material accumulation of nutrients from the soil can be ascertained and the analysis carried to any degree of refinement. There is no occasion to go into details here. Suffice it to say that such data give a record of the developmental history of the plants and permit of further analysis throwing light upon the major physiological functions.

It may be apposite to mention here the need for analytical work of this kind on plants in controlled environments; as I mentioned, development of the plant is modified to a quite remarkable degree by the level of external factors. At present there are no facilities available in this country for work of this kind; unfortunately they are very expensive to erect.

For the purpose of understanding the organization of the plant such studies of the mature organs will not suffice. The problem of growth and form is centred in the meristems or growing points in which cell multiplication takes place. Here the whole later symmetry of the plant is determined. It is in this connexion that the topic of leaf arrangement I mentioned earlier finds its place. The arrangement of leaves on stems, or flowers in such inflorescences as that of the sunflower, has been for some centuries the object of study by botanists. In succulent plants with reduced stems and crowded leaves, such as the house leek, or crowded inflorescences such as the sunflower, a remarkable symmetry appears in the form of intersecting spirals. The symmetrical arrangement of the mature organs, leaves or fruits, derives directly from the mode of origin of these appendages in the growing points. Unfortunately, these are of microscopic dimensions and thus open to study only after dissection and under high magnification. The problem of the determining factors leading to orderly geometrical arrangements is one of great fascination. My colleague Mr Richards has made a special study of the theoretical implications, and has utilized the data derived from the beautiful researches of Church, Snow and Wardlaw. He has been able to show that the whole development of the symmetry can be defined in terms of two growth characteristics of the growing point, namely, the relative growth rate of the meristem, and the intervals of time

separating the differentiation of successive lateral organs at the growing point—the so-called plastochron period. Both these physiological characteristics are open to precise investigation. The factors determining the positioning of the microscopic rudiments of organs, and this positioning is extremely accurate, are open to experimental investigation, as Snow and Wardlaw have shown. By modification of external factors the symmetry may be greatly modified. This problem, above all others, is capable of accurate mathematical treatment, as Richards has shown in the paper published in the S.E.B. Symposium on *Growth and Differentiation*. He has done much later work which I hope to communicate to this Society.

RELATIVE GROWTH AND FORM TRANSFORMATION

BY J. S. HUXLEY, F.R.S.

The transformation of form by quantitative distortion of a basic pattern is one aspect of the pattern of form change in general. The study of the relative growth of parts will illuminate this aspect of form change, but not others, such as the complete substitution of one pattern by another, as at metamorphosis in echinoderms, or during the process of amphibian gastrulation. It aims at providing understanding of the quantitative alterations that occur in form transformation by discovering the distribution of what for the present can only be called *growth potential* in the system; this in its turn will suggest lines of experimental attack to discover the physiological and biochemical bases of growth potential and its distribution.

A first step was to find out whether the growth of parts growing at a different rate from the body as a whole obeyed any general law. For differential growth of this kind, the term *allometry* has now been generally recognized, in place of the earlier *heterogony*; *isometry* is used of the special case when the organ grows at the same rate as the body.

The law of simple allometry appears to provide a first approximation to such a general law. It states that the size of a part, y, is related to that of some standard, x (whether the whole body, the rest of the body without y, or some part of the body selected as standard for reasons of convenience), according to the formula $y = bx^{\alpha}$, where b and α are constants.

The constant b, representing the value of y when $x = 1$, has no biological or general significance. The constant α, however, is of significance, since it can be considered as the ratio of the specific growth rates of y and x. These may and generally do vary with time (age); but so long as differential growth obeys the rule formulated above, their ratio remains constant. This will follow from the supposition: (a) that the growth involved is essentially multiplicative, depending on the progressive self-multiplication of some unit or units; (b) that different parts of the body differ in multiplicative rate, i.e. in specific growth rate; (c) that some factor operates to keep the ratio of specific growth rates between various parts a constant one.

There is thus at any period of development what may be called a *constant specific growth ratio between the growth of any two parts*. However, that this is not all is shown by the fact that the rule applies also to the final or stable size relations of part (*y*) and standard (*x*) (in, for example, secondary sexual size characters of male holometabolous insects, deer antlers, and limb regeneration in Crustacea), but *not* to their growth towards that stable relation. There is in fact what the French call an *allométrie de taille* as well as an *allométrie de croissance*—an allometric *size relation* as well as an allometric *process of growth*; and this can only be interpreted on the basis of some equilibrium which operates to regulate the relative growth rate of a part either up or down towards the attainment of a certain ratio with that of the body as a whole.

This is well demonstrated by the downward regulation of the growth of grafts of eyes from larger to smaller specimens of *Amblystoma*, and by grafts of eyes from one species of *Amblystoma* to another.* It may depend on the partition of some substance responsible for 'growth potential'. Both types of allometry can be subsumed under the general concept of an *equilibrium of specific growth ratio between the parts of the body*. This has been found to apply over a very wide range. It applies to limb segments of such extremely different construction and mode of growth as those of mammals and Crustacea; to combs of fowls and the nitrogen content of mealworms; to the antlers of deer and the horns of Goliath beetles.

The exponent α may conveniently be called the coefficient of relative growth, though in cases of allometric size relation it may be better styled the growth-partition coefficient.

Allometric growth may be either positive (α > 1) or negative (α < 1). When α = 1, we have isometric growth, but this, though often approximated to, is uncommon. Allometric growth is subject to various complications. The same organ may change its specific growth ratio during ontogeny (e.g. that of the large claw of male fiddler crabs of the genus *Uca* is lower after sexual maturity), or may have its growth equilibrium interfered with by increasing absolute size (e.g. gradual decrease of the coefficient α for the jaws of very large male stag beetles of various genera).

The process of allometric growth may start at different times in different organs, notably during early embryonic development; or moulting may introduce bimodality into the results (spider crabs, earwigs, etc. (Huxley, 1932, p. 68 f.)).

A general objection has been raised to ascribing any validity of principle to the allometric law, namely, that if the various parts of an organ (e.g. the segments of a limb) each obey the allometric formula in relation to some standard, it is impossible that their sum (e.g. the limb as a whole) should also obey it.

I do not believe that this is well founded. The large number of cases of quite different kinds of organs and other constituents of the body whose growth shows an approximation to the simple allometric formula indicates that some general principle must be involved, and it is difficult to see what this could be except a constant ratio between the rates of self-multiplication in different parts. I would

---

* The examples referred to in this paper are discussed in Huxley (1932) except where otherwise stated. This one is discussed on pp. 51 f., 191 f. A number of theoretical points are critically discussed in Reeve & Huxley (1945, p. 121).

suggest that this is the basic factor, but that there is some second-order factor at work which prevents the allometric formula from ever being precisely realized, and relates or adjusts the growth rates of separate parts to that of the whole. It is possible, for instance, that the increasing relative size of a part attained as the result of a high relative growth rate, might have a slight inhibitory effect on the relative growth rate itself.*

The next point which emerged in studies of allometric growth was the frequent existence of *growth gradients*, i.e. the regular graded distribution of growth potential (capacity for differential growth) within an organ or region of the body.

The most obvious examples are those in which a given limb, appendage, or well-defined region grows at different rates in the two sexes, especially when it is approximately isometric in one and strongly positively allometric in the other, as with the mandible of stag beetles, the abdomen of crabs, or the chelae of many decapod Crustacea. The most striking example of all is seen in fiddler crabs (*Uca*), in which *one* of the male chelae, like both the female chelae, is isometric, but the other strongly allometric.

In all such cases it can be seen that whereas the proportions of the parts of the allometric organ change with absolute size, and do so in a graded way, the region nearer the free end almost always showing the highest allometry, while the proximal region shows an allometry only slightly above that of the adjacent region of the body, those of its less allometric homologue do so to a much less extent, until in an isometric organ we find no change in the proportions of its parts with increase in its absolute size. In other words, the parts of the isometric organ are also isometric, while those of the organ which is allometric as a whole are themselves allometric relative to it, and their allometry is distributed in a regular way within the organ.

When the organ is made up of well-defined parts (segments of a limb, etc.) these give the most convenient units for measurement. But it is probable that the growth gradient is continuous within each segment (i.e. affects the organ as a whole, and not unit by unit). This is indicated by unpublished work of mine on lobster chelae, where the distances between spines on a single joint of the limb can be measured.

For most purposes, however, the growth gradient can be measured and represented by finding the relative growth coefficient of the separate segments of the limb (preferably relative to a standard in the rest of the body), and then plotting these as ordinates against the serial position of the segments as abscissa. This gives a graphical picture of the distribution of growth potential which cannot be obtained in any other way.

One and the same organ can change its growth gradient during development. Thus the limbs of sheep, both forelimb and hindlimb have an almost level gradient in the period just before birth, followed later (up to half-grown specimens) by a marked negative allometry of the limb as a whole, which then shows a marked

---

* There are many cases of second-order effects affecting simple laws. Boyle's law needs correction at high pressures on account of the size of the molecules themselves; Mendel's laws often need correction to allow for differential viability, etc.

growth gradient, decreasing distally (see Huxley, 1932, figure 49, based on a recalculation of Hammond's data).

The distribution of growth potential in the three planes of space can of course also be determined by measuring length, breadth and depth of a limb segment (see Huxley, 1932, figure 54).

It is clear that the existence of a growth gradient in an organ will cause a transformation of the form (proportions) of the organ. D'Arcy Thompson in his great book *Growth and form* developed a method for showing form transformations (of the same organ or organism at different stages, or as between homologous organs or organisms) by the use of Cartesian co-ordinates. Sometimes, as in his comparison of *Diodon* and the sunfish *Orthagoriscus*, it is clear that the (in this case evolutionary) transformation has affected the body in a differentially graded way; and many other indications of graded growth appear from his results.

Recently Medawar (1944, 1945) has extended D'Arcy Thompson's method, so that it is now possible to describe the two-dimensional form transformation of an organ or of the entire body mathematically, and to give a complete graphical representation of it.

This method, however, although providing much the most comprehensive picture or description of the process of form transformation, sheds little light on the possible biological causes or mechanisms underlying the process. On the other hand, analysis in terms of allometry does so, since it strongly suggests a graded distribution of some substance or process concerned with the regulation of relative growth rate.

Gradients based on calculation of relative growth coefficient should, I suggest, be called *allometric growth gradients*, to distinguish them from other types of growth gradients, such as gradients in change of absolute measurements. These latter can be readily detected by Medawar's method, but, I repeat, they do not appear to me to throw so much light on the possible biological basis of the facts as do allometric gradients. Professor Medawar tells me that the α values (relative growth coefficients) of different parts of a system undergoing form transformation can be obtained from his formulation, but that if an allometric growth gradient is required, it will generally be simpler to calculate it from measurements made directly for the purpose.

Allometric gradients are also of value for comparative purposes, as their heights and slopes can be quantitatively compared. Once the idea of growth gradients had arisen, it soon became clear that they operated within the body as a whole as well as within markedly allometric organs. By various methods, including those of allometry and of co-ordinate transformation, what may be called a growth profile of the whole body may be constructed. This is well brought out in regard to the growth of the limbs of higher Crustacea, plotted along the main axis of the body, when it is seen that the growth profile generally consists of a number of gradients, usually of low slope, running into each other. The different curvature of the two horns in two-horned rhinoceroses is evidence of the existence of an antero posterior gradient in the epidermal growth potential of the head of these animals (Huxley 1932, p. 150f.)

An interesting corollary of the general existence of growth gradients is that, in all animals with indeterminate growth, there is no fixed form, but form is a function of absolute size; and that apart from the rare cases of isometry, the precise proportions of, for example, different limbs are constantly changing. Their *precise* proportions, therefore, cannot be adaptively determined, but only their general proportions.

When accretionary growth occurs, as for instance in the shells of molluscs, a symmetrical linear growth gradient running down both margins of the shell-producing organ (mantle-edge) will result in the production of a shell in the form of a logarithmic spiral; when the growth gradient is asymmetrical on the two sides of the mantle edge, a turbinate spiral results. The horns of rhinoceroses are also logarithmic spirals, and for a similar reason. Once more, the interpretative value of the allometric hypothesis is demonstrated: a brilliant analysis and mathematical description, as provided by D'Arcy Thompson in *Growth and form*, is supplemented by a biological hypothesis, however, tentative it may be at this stage (see Huxley, 1932, chapter 5). A relatively simple mathematical analysis would be capable of revealing the quantitative gradients operating in different molluscan shells. This opens up an attractive and hitherto scarcely explored field of comparative study.

Perhaps the most unexpected result of allometric studies has been the demonstration that, in some cases at least, general growth gradients are deformable by alteration in the relative growth rate of single organs forming part of the general gradient. This is most readily demonstrated when, as in various decapod Crustacea, the male has much more strongly allometric chelae than the female. It then appears that the insertion of this region of high growth intensity into the general gradient slightly depresses the allometry of limbs immediately anterior to it, but slightly stimulates that of those immediately posterior. In other words, the general growth gradient is deformed in relation to the main axis of the body. We have no inkling as to what may be the biological basis for this phenomenon, but it points the way to certain types of experiment, as well as emphasizing the need for further measurements in crucial species (Huxley 1932, p. 120f.; see also pp. 115, 124 for indications of possible deformation of the growth profile of the body as a whole).

To sum up, both the analysis of the relative growth of single parts and the study of the form transformations of entire regions or of the body as a whole, have converged to indicate that, at any rate during the later phases of development, most animals show a graded distribution of the relative growth potential of their parts. Frequently, at least, this is resolvable into a series of allometric gradients of simple form; but there are indications that the separate gradients may interfere with each other, and that the distribution of growth potential in the entire body (growth profile) may be deformable as a whole.

Comparative studies show clearly that allometric gradients can be genetically controlled as units, and that both their height, their slope, and their size may be altered by natural selection (see Huxley 1932, pp. 46–48, 216f.).

Analysis is also indicating lines of experiment which would shed further light on the behaviour and possible biological basis of allometric gradients.

THE THEORY OF DIFFERENTIAL GROWTH ANALYSIS

By D. A. SHOLL, *Department of Anatomy, University College, London*

Growth will be defined as that attribute of living organisms which is manifested by a change of size of the individual, and we shall first consider growth formulae and curves. If y represents the size of any individual at time t, and if the growth process is regarded as continuous, we may define the absolute growth rate as dy/dt and the specific growth rate (which represents the change in y for change in time per unit amount of y) as $\dfrac{1}{y}\dfrac{dy}{dt}$.

Any statement which relates the size or growth rate of an organism to other variables will be called a growth formula; such statements are usually expressed in mathematical symbolism. Many growth formulae have been published and they have been derived by one of three methods.

The investigator who uses the first method usually makes the hypothesis that growth is a process which can be described *a priori* by means of a differential equation. Robertson used this method when he drew the analogy between growth and the autocatalytic chemical reaction.

The second method selects an arbitrary number of variables, such as the weight of the animal at birth, its basal metabolic rate, size of its hindlimb, and then makes an hypothesis about the relationship between growth and these factors. This procedure is also of an *a priori* type.

The third method derives a growth formula from an inspection of a plot of observations on squared paper and a curve is fitted to the pattern of these plotted points. It is clear that the best fitting curve will be a polynomial of sufficiently high order, but often some other type of function is chosen.

No matter which of these methods is used, a formula has been obtained, and to be of any value it must fit the observations reasonably well. 'Goodness of fit' must be tested by an objective method, for the fact that the graph of the curve derived from the formula appears, visually, to pass well through the points is no criterion; statistical tests are essential.

Moreover, it must be remembered that although a curve fits well, it does not follow that the formula from which it is derived is the unique description of the relationships between the variables and the theory upon which it is based consequently 'true'.

If the investigator has decided that a certain formula is of use, then it is essential that the parameters should be estimated by methods that are statistically efficient; indeed, not only must the parameters be efficiently estimated but the relevant standard errors must be determined. No tests of hypotheses and no comparisons are possible without such estimations. In addition, the assumptions upon which the methods of estimation depend must be borne in mind. For example, if a straight line is fitted to a set of data by the usual method of least squares, it is assumed that one of the variables, e.g. body weight, is known precisely, whereas the other, e.g.

organ weight, is normally distributed for each value of the first variable. It is not always remembered that it is consequently possible to fit two straight lines to any set of data assumed linear, and that these will only coincide if the correlation coefficient is unity. Many investigators seem to be unaware that a line may be fitted when both variables are subject to error; this method has been discussed by Richards & Kavanagh (1945).

The estimation of parameters when the formula is non-linear is more complicated. Sometimes it is possible to overcome the difficulties easily, but the precise fitting of such formulae as the Gompertz and logistic equations has not been possible until recently. The publication of Hartley's paper (1948) on the 'Estimation of non-linear parameters' has shown that the efficient estimation of the parameters in these and in similar cases is not only possible but does not involve complicated computation.

The danger of correlating indices is often overlooked. Over fifty years ago Karl Pearson (1897) examined this problem and pointed out *inter alia* that if a correlation coefficient were calculated between the index x/y and y, the spurious correlation could reach a value of 0·6.

Any discussion of the problem of relative growth must first examine the allometric formula

$$y = bx^a$$

This formula was, I think, first used many years ago by Snell (1891) in connexion with the relative sizes of brain and body weight. Dubois (1897) used it again during the nineties, but its use was widely extended by Huxley (1932).

Many observers plot their observations on a double logarithmic grid, and if the points appear to be linear, conclude that the allometric relationship holds. It must be remembered that a linear relationship will still give rise to a linear plot on a double logarithmic grid. Indeed, as D'Arcy Thompson suggested (1942, p. 207), the relationship between the sizes of the part and the whole organism is often linear. Figure 18 shows the results of plotting the length of certain parts of human beings against the height; it is clear that there are no indications of a departure from linearity.

The allometric formula contains two parameters; b denotes the size of the part when the value of x is unity and its value is clearly dependent on the units in which y and x are measured. The parameter a is the ratio of the two specific growth rates. The constancy of its value does not imply the constancy of these rates. Furthermore, the relationship supplies no information about growth or relative growth; time has, so to speak, been eliminated and the formula refers merely to relative size.

Very often the formula has been applied to data without the use of adequate statistical methods. Lines have been fitted by eye and parameters estimated inefficiently. When an attempt at precision has been made, the least squares method has been used, usually without consideration being given to the fact that if x is normally distributed, then the distribution of log x will not be normal. Moreover, if the plotted logarithms of the observations do not show linearity they have been split up into possibly linear segments, each having a 'constant differential growth ratio'. These dangerous and misleading methods have resulted in theories such as
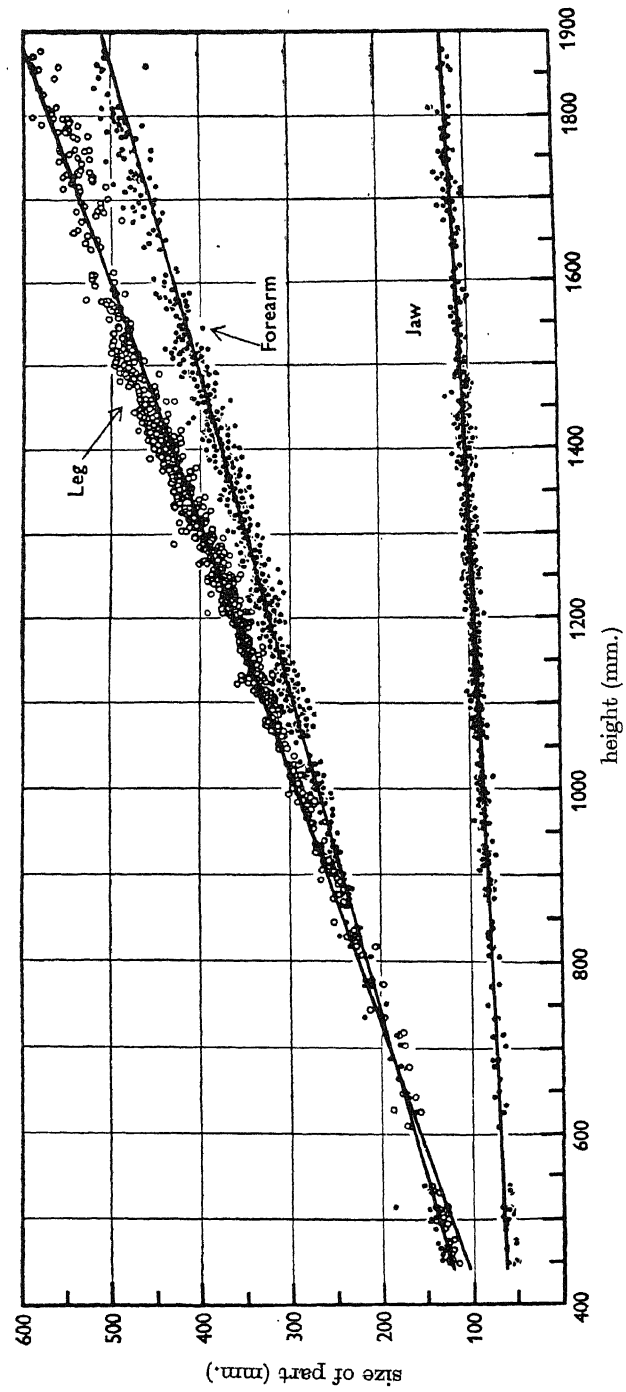
FIGURE 18. Measurements of limb and jaw sizes in man plotted against height. The fitted regression lines are also shown. (From Shepherd, Sholl & Vizoso (1949, p. 298.)

those sponsored by Dubois (1897), Lapicque (1898) and Brummelkamp (1939). This question has been fully discussed elsewhere (Sholl 1948).

A further difficulty inherent in all these methods is that data that are apparently of similar nature may demand very different treatments. Measurements may have been made on the same individual at different times; sometimes this is impossible, and the measurements may then comprise a series made on different animals at different ages. The statistical treatment of such different series demands careful consideration; these problems have been fully discussed by Richards & Kavanagh (1945) in *Essays on growth and form*.

The allometric approach has a further serious drawback. Only two parts can be compared at once; even then it might be preferable to plot the specific growth rate for each part and find the values of the ratio at different times. Certain alternative methods for the study of differential growth have been suggested but not widely used. The method of embedded co-ordinates employs a co-ordinate net applied to the individual with respect to selected reference points; the animal is photographed at different times and the deformation of the co-ordinates noted. This is a direct application of D'Arcy Thompson's method. One could possibly modify the method and use it in conjunction with the equipment used for somatotyping and photograph the individual with a standard grid projected on to its body; from the photographs it would be possible to calculate the deformation of the original grid which would be necessary to make it conform to the new proportions of the individual.

A further method which has been little explored may easily be understood by imagining small bits, e.g. patches of vital dye, embedded in the material and each patch labelled by suitable co-ordinates. After a lapse of time the co-ordinate values will have changed. From these pieces of information it is possible to derive simple formulae that enable the specific growth rates in volume, area and length to be found.

One of the fundamental needs for the adequate study of differential growth and, indeed, even the growth of an individual, is a long-term programme during which a number of individuals are measured regularly over a number of years. Adequate growth curves for a score of individuals would be of more value for investigating changes in shape and size than hundreds of curves founded on the mean values of different groups at different ages.

It is often maintained that the curve obtained from the measurements of size at different ages of different individuals gives information about the growth of some kind of hypothetical mean or ideal animal. This appears to be a vague concept, and any such curve may be very different from that for any individual. It is well known that there is a sudden spurt of growth in children at the time of puberty. Since the age at which puberty occurs is highly variable, the composite curve made from a number of individuals may fail to show this spurt although it is very clear in the curves of the individuals.

Methods of multivariate regression and discriminant analysis seem to offer promising techniques for the investigation of differential growth and, so far, have been little used for this purpose.

The dangers inherent in any method which depends upon a formula must be borne in mind. If the equation has been derived on *a priori* grounds, it may obscure factors of importance, although the very failure of the formula to meet the facts may lead to important discoveries. If the formula is empirical and results from fitting to the data, then it is merely a description of those data and extrapolation is perilous.

Finally, it is as essential for the experimental scientist to consider the assumptions and conditions upon which the validity of a statistical or mathematical method depends as for the engineer to remember the conditions for the validity of his formulae.

Transformation of shape

By P. B. Medawar, F.R.S., *Zoology Department, University of Birmingham*

We are lucky to be able to discuss to-day so simple a concept as that of 'growth'. Growth means change of size, and the size of an organism is something definite, unambiguous and measurable. It has been more than once suggested that 'growth' is an all but impenetrably obscure concept because it relates to a very heterogeneous process, i.e. because a great number of diverse factors contribute to change of size. One might equally well argue that the respiratory quotient was an ill-defined ratio because of the complexity of the respiratory process. The diversity of the causes of change of size does not imply that change of size is an obscure idea. As biological concepts go, it is an uncommonly clear one. It can, of course, be made as obscure as we wish by arbitrary subdivisions of its terms of reference. One may distinguish, for example, between a 'true' sort of growth and a 'mere' sort of growth; molecular replication is 'true', and the intussusception of water is 'mere'. This sort of distinction is a surviving remnant of the habit of mind that distinguishes between the protoplasm of the animal, which is truly alive, and the rest of it, which is dead. There are, indeed, profound differences between the processes that severally contribute to change of size, but they are not differences of which terms like 'true' and 'mere' can usefully be applied.

A *description* of growth must necessarily precede an investigation of growth processes, just as an analysis of the mechanism of nerve regeneration or skin healing cannot be begun unless we know what is actually going on, in an anatomical sense, when nerves regenerate and skin heals. The recording of growth is thus *primarily* an exercise in anatomy or descriptive embryology. Because size is measured numerically, the description of growth is necessarily mathematical in the everyday sense, but this is not inconsistent with its being anatomical as well. An anatomist who says that one organism is ever so much larger than another, or is longer and thinner than another, is making statements which in a general sense are mathematical. They are mathematical statements and they are mathematically imprecise. There are some purposes for which documentary information about growth is best presented in the form of parallel columns of figures for sizes and ages, or weights and heights, as it is in some weighing machines on railway plat-

forms. A person who weighs himself is interested in his own weight and may be interested in the mean weight of other people of his own age or height. One pair of entries among a double column of figures can provide him with this information instantly, and it would be unkind to expect him to extract it from a growth equation intended to be applicable to all ages. But for most purposes the construction of growth equations is very well worth while: they expedite the analysis of growth rates; they allow comparison between the modes of growth of related organisms; and, most important, they make it possible to frame inductive 'laws' of growth after an investigation of the general analytic properties which the growth equations of different organisms share in common (Medawar 1945).

It may be worth mentioning two among several types of hazards and difficulties that beset the construction and use of growth equations. One is the problem of disentangling curves of growth from curves of size frequency distribution. Suppose we follow 100 organisms from birth until death and weigh them at intervals. If we plot the sum of their weights against age we describe a curve which represents the growth in mass of the population as a whole, the growth of its 'biomass'. If we divide each summed weight by 100 and plot the quotient against age, we are in fact plotting the mean contribution of an organism to the biomass of the population. The change in the mean contribution of an organism to the biomass of the population is itself a perfectly definite and respectable measurement, but difficulties arise when it is equated, as it so often is, to the growth curve of the individual. The two may differ profoundly, particularly when growth is episodic rather than more or less smoothly continuous. The accompanying figures illustrate by a deliberately extreme example how gross the disparity may be. Figure 19 illustrates a hypothetical organism which undergoes a sudden increment of size within the period of measurement. A group of other such organisms undergo exactly similar spurts, save that they are not synchronous; we may suppose that the times t at which the spurts take place are normally distributed about a mean $\bar{t}$ with unit variance (figure 20).

Figure 21 is computed from the mean sizes of the organisms at each age. The curve so figured accurately and properly describes the mean contribution of an organism to the total biomass; it does not describe, and is, indeed, totally unlike, the growth curve of an individual (figure 19). Its sigmoid segment, to which one might so easily be tempted to fit a logistic equation, is a tilted frequency distribution curve. This is, admittedly, an extreme example, but the growth of some organisms (e.g. of insects over a period that includes a moult) is not grotesquely far removed from it.

A second problem is concerned with what may be called the 'fallacy of summation'. Let it be supposed that $y = bx^{\alpha_1}$ is a theoretically valid description of the relationship between the size y of a certain organ and the size x of the organism as a whole during its growth; and that $z = cx^{\alpha_2}$ in the same way validly describes the relation to the size of the whole organism of an organ of weight z. For example, y might stand for the length of the leg, z for the length of the thigh, and x for total height; and it is being suggested that the length of the thigh and the length of the leg increase in proportion to constant powers (α_1 and α_2 respectively) of the height

of the body. If this is so, the length of the limb as a whole, $y+z$, cannot itself increase in proportion to a constant power of body size, i.e. it is not in general true that one can find values of d and α_3 such that $(y+z)=dx^{\alpha_3}$. The sums and averages of these power functions are not functions *of the same sort*. Of course, it could still be true that an equation of this form was an adequate *empirical* description of the growth of the whole limb; but it cannot be a theoretically valid description of the growth of the limb as a whole if it is a theoretically valid description of the growth of its two constituent parts. Difficulties of this sort, first pointed out by J. B. S. Haldane, cease to arise so soon as it is admitted that growth equations are purely empirical in standing.
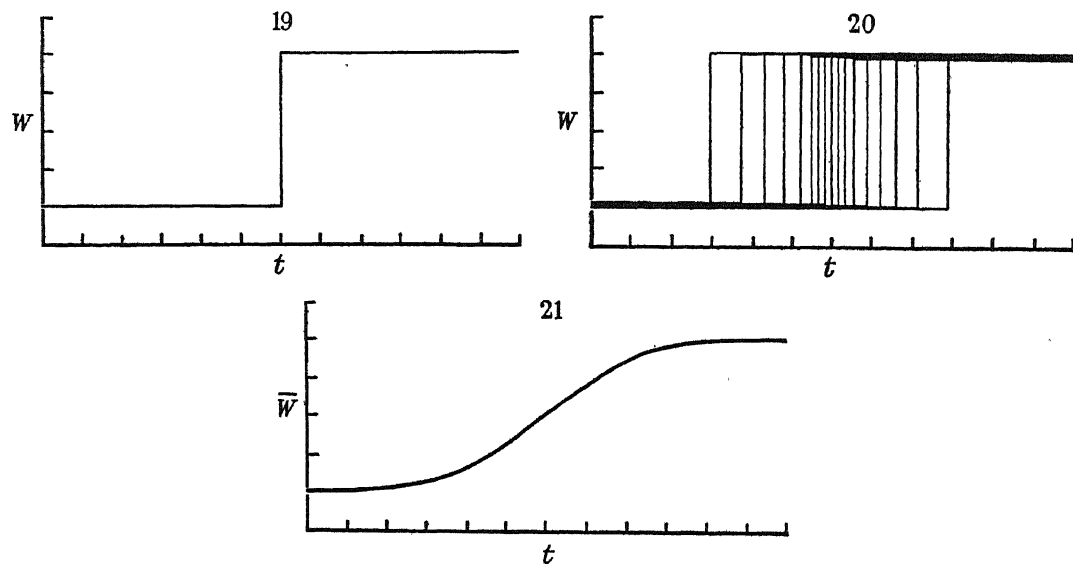
FIGURES 19 to 21.

The equation that has just been discussed is Huxley's well-known and widely used equation for allometric or differential growth. In its integral form, $y=bx^{\alpha}$, it asserts that a part of an organism will generally grow in proportion to some constant power of the size of the whole. In so far as its status is empirical, one cannot ask whether it is right or wrong, but only whether or not it is expedient and useful. Its rationale is this. The growth of the organism is fundamentally of the multiplicative type; and it is this fact that makes it expedient to use the specific growth rate, rather than the simple growth rate, when describing the course of growth as a function of age. This being so, then whatever merit attaches to the use of the specific growth rate as a representation of growth in time should justify the use of a *ratio* of specific growth rates to describe the differential growth of two parts. The specific growth rate of a part y is $\dfrac{1}{y}\dfrac{dy}{dt}$ or $\dfrac{d(\log_e y)}{dt}$ and of a part x is $\dfrac{1}{x}\dfrac{dx}{dt}=\dfrac{d(\log_e x)}{dt}$, and their ratio is

$$\frac{x\,dy}{y\,dx} \quad \text{or} \quad \frac{d\log y}{d\log x}.$$

In the simplest case this ratio is constant, and if we equate it to α it follows, by integration, that $y = bx^\alpha$, where b is a constant of integration. It is sometimes suggested that Huxley's mode of formulation is incorrect because α can hardly be expected in the general case to be constant. This would seem to imply that the use of the specific growth function is admissible only when the specific growth rate is constant, i.e. only when it is being used to describe the properties of a system increasing by continuous compound interest. In practice, however, it is used with satisfaction to describe growth systems of which the specific growth rates are by no means constant. In like manner the ratio of two specific growth rates is an acceptable formulation of differential growth rate even if their ratio is inconstant. (It has the additional merit that simply proportionate growth—mere enlargement— is representable as the special case $\alpha = 1$.) I do not think that Huxley's formulation can be dismissed because some sorts of biological growth are accretionary and not multiplicative. It is sometimes forgotten that many sorts of accretionary growth are in fact subsidized by multiplicative growth systems. The growth of nails is accretionary; but nails are made of the concrescent squamous scales of epidermal cells, and epidermal cells constitute a multiplicative growth system, since some of the growth products of epidermal cells are themselves capable of growing. The nail is a by-product of the multiplicative activity of the nail-forming epithelium. The growth of the mineral ingredients of bone is also subsidized by a multiplicative growth system, though in a different way. In short, to say that the growth of organisms is fundamentally of the multiplicative type is an acceptable generalization, and Huxley's formulation of differential growth gives this generalization the credit that is its due.

The piecemeal analysis of differential growth is not of course a complete description of change of shape, though it describes certain ingredients of a process of transformation. There are certain purposes for which it is useful to have a comprehensive picture of change of shape, and I should like to say a few words about it.

Transformation, or change of shape, is no more complex a notion than growth, or change of size; nor is shape itself any more obscure than size. It is true that shape cannot be measured, because it is a complex quantity. No organism will ever be said to be, or to have been, $2\frac{1}{2}$ Thompsons in shape. But shapes are always definable in principle: that is, I can write... 'Circle $(x, y) = 0$' or 'Face $(x, y) = 0$' or '*Diodon* $(x, y) = 0$' to affirm that the variables x and y are to take such pairs of values as will describe the outline of a circle or of a plane drawing of a face or of the fish *Diodon*. Shapes are geometrically defined by *sets* of points, namely, the sets whose members satisfy functional relationships of the sort that have just been attributed to them.

The principle that is involved in describing changes of shape is a very simple one. A change of shape is defined by a change of variables. If I substitute a new variable x' for the old variable x in accordance with the rule $x' = 2x$, then Circle $(x, y) = 0$ becomes transformed into Circle $(\frac{1}{2}x', y) = 0$, and the set of points (x', y) describes the outline of an ellipse. Other such substitutions $x'' = 3x$, $x''' = 4x$, each governed by different substitution rules of the same family, define outlines that become successively more elliptical. In this case, the general rule of transformation

is clearly $x^* = kx$, where x^* stands for the new variable that is to replace the old, and k is a variable parameter, i.e. a variable quantity that takes a constant value in each particular context. If I wished to represent the deformation of a circle into outlines that became progressively and continuously more elliptical in the course of time, then it would be necessary to discover the relationship between the variable parameter k and the age of the system t. When such a relationship $k = f(t)$ has been established, then, knowing t, the value of k can be determined; knowing k, the particular substitution rule is given, and so in turn the particular transformation. The continuous variation of t (supposing $f(t)$ to be a continuous function) would then determine a continuous set of transformations of the circle in time. Representing the changing shape of the circle as a continuous one-parameter transformation scheme is exactly analagous to representing size as a continuous function of age. Clearly the scheme does not depend upon one's being able to define the shape of a circle, face or fish in concrete algebraical terms; it is change of shape, not shape, that is under investigation, and the same transformation scheme may be applied indifferently to them all. In like manner we are not obliged to attach particular concrete values to a simple scalar quantity undergoing variation in time when we are interested only in its mode of variation; the rate of interest offered by a savings bank applies indifferently to pounds, shillings and pence, and sixpence invested grows at the same specific rate as six pounds.

The 'Method of transformations' was, of course, devised by D'Arcy Thompson (1916). Its recent development (Medawar 1944, 1945; Needham 1950) has enlarged upon it in two ways: the method has been rendered analytic, and not merely geometrical, and it has been made competent to describe *continuous* changes of shape, i.e. to substitute a cinema performance for D'Arcy Thompson's lantern slides.

In the study of growth in time, it may well happen that the mode of growth changes during a lifetime, so that functions of different sorts must be used to describe its several phases—first a compound interest function, maybe, and later a function describing a die-away of growth rate. The analogous state of affairs when dealing with continuous one-parameter transformations is to change the type of function which provides the rule for replacing one variable by another. In the example that has been considered above, the same function $x^* = kx$ governed each substitution, and one substitution differed from another simply because a different value of the parameter k was allotted to it. One of the first things to be decided on in analyzing a change of shape is, then, whether or not one transformation function is competent to describe a continuous change of shape.

Another analogy may be pressed. Having devised an equation which describes size as a function of age, we may then apply the ordinary procedures of the calculus to work out growth rates and accelerations of growth. Methods have been described by which the same *sort* of information may be extracted from equations defining continuous sets of transformations; one can answer such questions as, 'Does an organism increase in shape as it develops?' or 'Is the rate of change of shape uniform throughout life?' It will be clear from what I have said that the method of continuous transformations contains *all* the information that may be offered by a piecemeal

analysis such as that made use of by Huxley. This does not mean that it is a 'better' method. A method is good or bad as it presents more or less lucidly and directly information of the type which we want to extract from the data. If we were specifically interested in the relationship between leg length and body length, say to provide information relevant to a study of nerve regeneration, or to guide a manufacturer in the design of armchairs or trousers, or because we were concerned with the hydrostatic shortcomings of the vascular supply of the feet of very tall men, then it would clearly be unnecessary, and therefore a waste of time, to undertake a complete Thompsonian analysis.

I should like to conclude by calling attention to a very important shortcoming of the 'method of transformations'—or, rather, to a class of problems which might be expected to lie within its ambit of competence, but which in fact does not. The functions that determine the rules by which old variables are replaced by new are allowed a wide range of latitude, but beyond one limit they may not trespass: the functions must be continuous and must be single-valued both ways, i.e. must establish a one-to-one correspondence between the points of the new figure and the old. The effect of this restriction is to confine the use of the method to the comparison of forms which are of the same *order* of complexity, i.e. roughly speaking to forms which are derivable one from another by plastic deformation, however grotesque or extreme it may be. In technical language, the forms that are to be compared must be homeomorphic. It is unfortunately true that many of the most interesting transformations of development *do* involve changes in order of complexity: a frog's embryo that has produced a new closed body space by the splitting of a formerly single-layered sheet of mesoderm has increased in order of complexity. Such changes are outside the competence of the method of transformations. From time to time new methods for describing change of complexity have been suggested, but they do appear to have had a fruitful outcome.

THE PLACE OF STATISTICS IN THE STUDY OF GROWTH AND FORM

By F. YATES, F.R.S., *Rothamsted Experimental Station, Harpenden, Herts*

Much of the statistical work involved in the study of growth and form consists of the fitting of some mathematical relation to the observed data. This fitting may be carried out with a number of rather different purposes in view, which may be classified as follows:

(a) descriptive, as, for example, when a table of values is replaced by a simple equation;

(b) for the purpose of smoothing data;

(c) to enable objective comparisons to be made between groups of data, similar curves being fitted to the different groups;

(d) to test whether some mathematical relation derived from other considerations is confirmed by the data;

(e) in an attempt to obtain an indication of the underlying physical laws.

These objectives are not mutually exclusive, but the types of relation and the statistical procedures that are appropriate differ somewhat in the different cases.

Three different steps are involved in the fitting process. The first consists of the choice of the type of relation, the second is the estimation of the values of the parameters which enter into the equation representing the chosen relation, and the third is the testing of the agreement of the fitted curve with the original data.

In cases (d) or (e) above, the choice of relation is not a statistical matter—if a relation given a priori requires to be tested this must be fitted, if an indication of the underlying physical law is required relations which can bear a physical interpretation must be chosen. Thus a polynomial curve, though mathematically simple, is not usually capable of any simple physical interpretation. At best it can be regarded as an approximation, over the range of the observations, to some more complicated relation.

In cases (a), (b) and (c), on the other hand, the choice is governed by convenience of fitting, and by the need for a relation which fits the data with sufficient accuracy for the purpose in hand. These are statistical matters. The question of convenience of fitting is an important one statistically because a convenient method will be used far more extensively than will a method which requires elaborate and tedious calculations. It is one of the advantages of the allometric equation that it is easily fitted by transformation of both sets of measurements to logarithms.

Once the type of curve has been decided, the efficient estimation of the values of the parameter, and the testing of the goodness of fit, can be dealt with by known statistical theory, though the fitting of many types of curve presents difficulties both of theory and of computational procedure. If we have knowledge of, or hypotheses concerning, the underlying physical laws, the values of the parameters determined by the fitting will provide estimates of the undetermined physical constants, and the tests of goodness of fit can be taken as tests of the correctness of the assumed laws.

Adequacy of fit does not, however, give any definite evidence that the underlying physical laws are of the type represented by the relation. This is obvious in the case of empirical relations, such as polynomials, but is true also of relations based on physical laws. Such laws are only confirmed within the limits of error of the observations. Inaccurate or fragmentary observations are often taken as confirmation of a law from which more accurate and extensive data subsequently reveal significant deviations.

For comparative purposes, such as are required when the growth and form of different species have to be contrasted, and in experimental work, empirical relations are usually satisfactory, even when the fit is by no means perfect. Provided the same form of equation is fitted to different animals or groups which it is desired shall be compared, failure to fit perfectly may well be of little consequence, since the resulting inaccuracies affect each of the groups similarly. Simple types of curve, such as polynomials, which are easy to fit, are of great value in this type of work. In some sampling measurements on the growth of wheat, for example, which were carried out on replicated plots sown to a pair of varieties at different centres over a number of years, the maximum growth rate

and its date of occurrence were determined by fitting cubic polynomials to the shoot heights. Comparison of the original data with the fitted curves showed that the determination of the point of maximum growth rate by this method was somewhat biased, owing to more marked curvature of the growth curve above the point of inflexion than below. Nevertheless, this simple and convenient method of determining the date and rate of maximum growth gave results which were quite adequate for comparative purposes.

TABLE 1. MEASUREMENTS, IN ARBITRARY UNITS OF VARIOUS FEATURES OF A
HUMAN BEING FROM THE 5TH MONTH OF FOETAL LIFE TO MATURITY

	time in years from conception					
	0·42	0·75	2·75	6·75	12·75	25·75
base-line	0·00	0·00	0·00	0·00	0·00	0·00
fork	1·00	1·59	1·80	1·94	2·09	2·23
navel	1·81	2·29	2·50	2·67	2·77	2·90
nipples	2·52	3·00	3·20	3·38	3·51	3·64
chin	3·20	3·62	3·83	4·03	4·13	4·22
total height	4·85	4·85	4·85	4·85	4·85	4·85

TABLE 2. DIFFERENCES FROM TABLE 1 EXPRESSED AS PERCENTAGES
OF THE TOTAL HEIGHT

	time in years from conception					
	0·42	0·75	2·75	6·75	12·75	25·75
foot to fork	20·6	32·8	37·1	40·0	43·0	46·0
fork to navel	16·7	14·4	14·4	15·0	14·0	13·8
navel to nipples	14·6	14·6	14·4	14·6	15·3	15·3
nipples to chin	14·0	12·8	13·0	13·4	12·8	12·0
chin to top of head	34·0	25·4	21·0	16·9	14·8	13·0
	99·9	100·0	99·9	99·9	99·9	100·1

On the other hand, the methods used must be such that no relevant distortion of the data results from the process of fitting. Methods which are suitable for some purposes are quite unsuitable for others. In particular, there is a danger that a spurious regularity may be introduced into the data by the process of fitting. Table 1 is reproduced from data given by Medawar (1944, 1945). The original data are due to Jackson (1915).* These data were used by Medawar to demonstrate the statistical methods appropriate to the quantitative analysis of growth gradients, and the conclusion was reached that a gradient was exhibited by the data. As presented, the data are not in the form most suitable for inspection, and I have therefore transformed them by taking the differences and expressing these as percentages of the nominal total height. These are shown in table 2, the first line

* I should perhaps make it clear that this example was chosen for illustrative purposes only, and it is not intended as a criticism of the methods developed by Professor Medawar for providing quantitative statements of growth gradients. Nor is it intended as a disproof of the existence of a simple gradient in human beings, any more than was Professor Medawar's original examination intended as a proof of such a gradient. To settle this point the more accurate and detailed data now available should be examined.

of which, for example, indicates the percentage of the total height contributed by the legs (from foot to fork) at various ages.

It is immediately clear from inspection of these percentages that the main relative growth features are the steadily increasing proportion of the total length contributed by the legs, and the decreasing proportion contributed by the head. Subsequent to birth there is very little change in the other three percentages, but during the pre-natal period there is a fall in the fork-to-navel percentage. If a growth gradient existed the change in this percentage should be intermediate between the changes in the first and third percentages, i.e. it should show a rise. As far as the trunk is concerned, therefore, the gradient hypothesis is not supported by the data, which in fact indicate that the whole of the trunk grows at approximately the same rate. The gradient was 'read into' the data by the fitting of polynomials to the joint plots of column 1 of table 1 (x co-ordinates) against each of the other columns in turn (y co-ordinates). In his original paper Medawar himself observes that the points representing the fork show consistent deviations from these polynomials.

Another example of distortion of the data resulting from faulty statistical processes, which has already been mentioned by Medawar in the present discussion, is that which results from taking the mean of the growth curves of a number of individuals. Such a mean growth curve is perfectly adequate for determining the size distribution of individuals at successive ages, or the mean growth curve of the population in the aggregate, and similar purposes. For such purposes, provided the average growth curve is not changing with time, it is immaterial whether the data relate to successive measurements on the same group of individuals, or to simultaneous measurements on groups of individuals of differing ages. The important requirement is that the sample shall be sufficiently large to determine the required means with adequate accuracy, either directly or after smoothing the data by the fitting of suitable curves.

If, on the other hand, the growth of individuals requires to be studied such mean curves are entirely misleading. Thus if there is a sudden change in growth rate, say at the onset of puberty, and this onset occurs at different ages in different individuals, a growth curve constructed by averaging all individuals of a given age will not show a sudden change, but only a smooth transition from one growth rate to the other. The sudden change and its different ages of onset can best be brought out by the study of growth curves of different individuals. For this we require consecutive observations on the same individuals. Individual growth curves can then be fitted and the characteristics of these curves summarized.

If such observations are impossible little can be done unless some indication of the stage of growth reached by the individual is available. If, for example, the onset of puberty can be determined by some physical characteristic, the data can be classified according to this characteristic.

The purposes for which the fitted curves are required not only affect the choice of curve, but also the process of fitting. In particular, it is important to see that, when the fitting is used to estimate the values of undetermined physical constants, biases are not introduced into the estimated parameters. The case of a linear

regression when both sets of observations are subject to error is a well-known elementary example of this. If the true law is

$$y = \alpha + \beta x,$$

then the customary estimate b of β

$$b = \frac{S(x - \bar{x})(y - \bar{y})}{S(x - \bar{x})^2}$$

is correct for predicting values of y for further individuals of which only the values of x are observed, but the estimate b' of β which is virtually unbiased for large samples is given by

$$b' = (1 + \gamma)b, \tag{1}$$

where γ is the ratio of the error variance to the true variance of x.

The degree of efficiency required in the fitting depends on the objects in view. If tests of goodness of fit are required then methods of fitting must be reasonably efficient, as if this is not so, the tests will be vitiated. On the other hand, for many purposes methods which are not fully efficient (including graphical methods) will be quite adequate. The danger with graphical methods is that they are to a certain extent subjective, and are therefore liable to give misleading results in inexperienced hands.

The efficiency of a fitting process cannot be properly considered without knowledge of the distribution of those parts of the variation that can be regarded as equivalent to error. With the ordinary types of fitted curve, particularly regressions, used in the study of growth and form, this appears to be a question of more theoretical than practical importance as far as estimates of the parameter are concerned. On the other hand, certain of the tests of goodness of fit may be vitiated if the data are markedly non-normal.

Quite apart from questions of efficiency of fitting it is important to see that the parameters of which the estimates are presented are so chosen that these estimates are not seriously influenced by random errors in other parameters. In fitting the linear regression equation

$$y = \alpha + \beta x,$$

for example, the estimate a of α is given by

$$a = \bar{y} - b\bar{x},$$

which, if $\bar{x}$ is at all large, is subject to considerable errors due to errors in the estimate b of β. If, instead, we use the equivalent equation

$$y = \alpha' + \beta(x - x_0),$$

where x_0 is a 'standard' value of x in the neighbourhood of $\bar{x}$, errors in the estimate a' of α' given by

$$a' = \bar{y} - b(\bar{x} - x_0)$$

due to errors in b are small. For purposes of comparing regressions derived from different groups of material with similar but not identical $\bar{x}$, the second form of the expression is much to be preferred.

This point arises also in the allometric equation, which involves a linear regression of $\log y$ on $\log x$. Here the standardized form

$$y = k'x^b/x_0^b$$

should be used.

We may at this juncture briefly consider the statistical processes that are appropriate to the handling of simultaneous measurements of a number of characteristics. Classical examples of this type of problem are provided by the extensive series of skull measurements collected in anthropometric studies, and the measurements of body shape and size on living populations of human beings.

For descriptive purposes such measurements can· be summarized by means of frequency distributions of the original measurements. A further stage in condensation is to calculate various characteristics of these distributions, such as the means, variances, covariances and correlations between different variates. If, however, the number of variates is at all large, a frequency distribution of all variates simultaneously becomes very complex, and little or nothing is gained by attempting to present such a distribution in place of the original data. For many purposes the joint frequency distributions of all pairs of variates provide all the necessary information. Alternatively, a compromise may be made by picking out the most important variate and giving joint frequency distributions of the other variates for each separate class of this variate.

If it is desired to express one character in terms of one or more of the other characters regression analysis may be used. This provides a useful descriptive method when some one character such as age is regarded as of outstanding importance. There are many variants of the method. We may, for example, consider the relation of height and weight after eliminating the effect of age by calculating the regression of weight on height for each age group separately, or taking a single partial regression with age eliminated.

Clearly regression equations of this type cannot be regarded as having any great physical significance. The existence of a pair of regression coefficients between similar variates such as height and weight is itself an indication that such coefficients cannot be regarded as estimates of any underlying physical quantities. In pairs of variates such as height and age, where age can be regarded as influencing height, whereas height cannot influence age, the regression of height on age is of more physical interest than that of age on height, but even here, in the absence of any biological theory of how age influences height, the regression equation must be regarded as primarily descriptive.

The duality of regression coefficients between two similar variates may be overcome in various ways. If, apart from errors of observation, or sources of variation which may be regarded as equivalent to errors, the relation between two variates is truly linear, we can, by the method given above, make an estimate which is unbiased in large samples of the coefficient β if we know the error variance of x, or of its reciprocal, if we know the error variance of y. From equation (1) above these estimates are

$$b' = (1 + \gamma_x) \frac{\operatorname{cov} xy}{V(x)},$$

$$\frac{1}{b'} = (1 + \gamma_y) \frac{\operatorname{cov} xy}{V(y)}.$$

If neither error variance can be estimated directly, we may be prepared under certain circumstances, when x and y are similar variates, to assume that the ratios

of the error variances to the true variances are equal, i.e. that $\gamma_x = \gamma_y$. In this case we have the combined estimate

$$b' = \sqrt{\frac{V(y)}{V(x)}} = \sqrt{\frac{b_{y.x}}{b_{x.y}}}. \tag{2}$$

It would be foolish to use this estimate, however, unless γ_x and γ_y are both small, which will only occur if the correlation coefficient is nearly ± 1. The two regression lines will then be nearly coincident, with $b_{y.x}$ nearly equal to $1/b_{x.y}$.

Alternatively, a line may be chosen such that the sum of the squares of the perpendicular distances from this line of the points representing the observations is minimum. This is, in fact, the line passing through the centroid of these points whose slope is one of the roots of the equation

$$\lambda^2 \operatorname{cov} xy + \lambda\{V(x) - V(y)\} - \operatorname{cov} xy = 0.$$

This line is coincident with that given by equation (2) if $r = \pm 1$, or $V(x) = V(y)$.

A measure of the closeness of association between two variates is provided by the correlation coefficient, and the effects of other variates can be eliminated by the use of partial correlation. Correlation analysis, however, has proved to be of limited use, and is generally less informative than regression analysis.

In addition to its descriptive function a regression equation also has the much more concrete function of providing an equation by means of which knowledge of one or more characters can be used to predict the most likely value of another character. The success of the prediction depends on the closeness of the relation between the different characters. There must, of course, be a basic set of measurements in which both characters are measured and from which the relationship can be established. Equally important, we must be confident that these measurements are carried out on material comparable to that on which a prediction is required. Thus, if the body form of human beings is established by a series of carefully conducted measurements on a more or less random sample, the distribution of these measurements on further groups of human beings can be predicted from key measurements such as weight and height, provided we are satisfied that the new group does not differ appreciably from the original population in the relation to weight and height of the measurements in which we are interested, even though the actual distribution of weight and height in the subgroup is somewhat different. Suppositions of this kind are clearly unlikely to be absolutely true, but may be sufficiently close to the truth for practical purposes.

The intrinsic relationships between a set of multiple measurements, in so far as they are linear, can be embodied in the matrix of variances and covariances, from which the correlation matrix can be immediately deduced. A correlation matrix, however, is singularly uninformative, and various forms of multivariate analysis have been devised in an endeavour to determine the influences that give rise to the observed correlations. The general principle is to see if the measurements can be represented as linear functions of a set of factors, each factor having a different value for each individual. Obviously this form of representation is only of interest if the number of factors is substantially less than the number of measurements.

Factor analysis has suffered from an accretion of approximate methods. The most rigorous form of analysis is the resolution into principal components. These components are given by the roots of λ in the equation

$$
\begin{vmatrix}
1-\lambda & r_{12} & r_{13} & \cdots \\
r_{12} & 1-\lambda & r_{23} & \cdots \\
r_{13} & r_{23} & 1-\lambda & \cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{vmatrix} = 0,
$$

where r_{st} is the correlation between the measurements s and t. The magnitude of the respective roots indicates the amount of the variance accounted for by each.

Unfortunately there seems to be little ground for believing that the factors which are determined by an analysis of this type correspond to any real physical entities. The principal components are orthogonal in the sense that they are uncorrelated. There is, however, no reason to imagine that the physical entities which give rise to them are uncorrelated, and consequently it must not be expected if, for example, the data are accounted for by two principal components, that these components themselves represent real physical entities. The real physical entities, if they exist, may be any pair of linear functions of these components. The most that can be said from the fact that a complicated correlation table can be accounted for by, say, a couple of principal components, is that there is evidence that the measurements in question are primarily influenced by only two variates. Here again, however, the conclusion is tentative. More accurate or extended observations may well reveal further complexities.

Discriminant function analysis, though somewhat similar in form to principal component analysis, has in fact very different functions. If individuals on which multiple measurements have been made can be divided into groups, it may be of interest to ask ourselves in what ways these groups differ. This question can in part be answered by examining each measurement separately, but for purposes of condensation and of assigning further unclassified members of the population to one or other of the groups a combined function of all the measurements is sometimes required. We can, in fact, determine the linear function of all the measurements which differs most between group and group relative to its variation within groups. This function is called the discriminant function. With three or more groups further linear functions can be isolated in a similar manner to the different principal components in component analysis. We can thus ascertain whether the differences between the groups can be described in terms of variations in a single linear function, or whether two or more such functions are required.

Multivariate analysis is still in its infancy. A considerable body of theory is now in existence, but apart from the simpler forms of regression analysis practical applications have lagged behind the theory. The reason for this, I think, is the lack of calculating machines which will deal conveniently with this type of problem, in the way that the ordinary desk calculator deals with problems involving a single variate. Without contact with practice theory soon becomes sterile. The very striking advances in the statistical methodology appropriate to a single variate,

governing the fields of analysis of variance, tests of significance, sampling errors, design of experiments etc., appear to be attributable not simply to the development of the relevant theory, but to the continued interaction between theory and practice that has resulted from the routine use of desk calculators of the modern type. There seems every reason to expect that the modern developments in punched card machines and electronic calculators will lead to similar progress in the multivariate field.

At the beginning of this paper, I stated that one of the purposes in the fitting of mathematical relations was to obtain an indication of the underlying physical laws. In concluding, I should perhaps make it clear that I do not personally consider that this line of attack is likely to prove very fruitful, owing to the complexity of the phenomena involved, and particularly owing to genetical and evolutionary considerations. From what we now know of genetics and evolution it seems certain that the size and form of the adults of many species are very closely controlled by selective influences. It is characteristic of such influences that when a change results in a selective advantage, any means by which this change can be brought about, however apparently complicated, may become incorporated in the species. In other words, in such species the actual controls governing the final size and form are likely to be a complicated jumble of detailed controls superimposed on one another, conditioned in large part by historic evolutionary development, rather than a simple and orderly set of controls such as might be arrived at by designing the organism completely afresh to a given specification. The embryonic development of human beings, for example, provides striking illustration of how persistent these historical factors can be.

Selective influences will, of course, also operate on growing organisms, but for many species the total period spent in any one stage of growth is very considerably less than that spent in the adult form. A greater or less degree of protection is also often extended to the young by their parents. Consequently in such species selective influences on the young are likely to be less potent, and we may expect to find less perfect balance between the different parts of an organism during the development stages than exists in the final adult.

The form of the newly born organism, moreover, is likely to be very different from that of the adult, since the size and shape of the different parts of the organism must be such that development in the confined space of the egg or the womb is possible. The proportions necessary for survival in the very young may also be very different from those in the adult form.

In species in which the young are protected, therefore, growth may be looked on as a developmental process of which the most important characteristic, from the point of view of survival, is that a given adult form shall be reached. The transition from the initial form of the newly born organism to the adult form may be governed by the somewhat general control mechanisms that are implicit in allometric relations and growth gradients, but other types of control mechanism which give the same end-results and do not lead to too violent disharmonies of form during growth will be equally satisfactory from the functional point of view. The data quoted above for human beings indicate such an alternative type of

control mechanism where the whole of the trunk appears to be growing isometrically at a different rate from the head, and from the legs. Thus the head, trunk and legs may be allometrically related (or approximately so), but without a smooth growth gradient extending over the whole body.

Another example of the complex nature of the genetical controls of growth and form is provided by Waddington's account in the present discussion of his studies of the growth of wings in *Drosophila*.

It seems to me improbable that any purely statistical approach will throw much light on the underlying laws and control mechanisms that govern such complex phenomena. The proper role of statistics appears to be the much humbler but still vital one of isolating the salient features of the multiple measurements that represent growth and form, so that their interrelations with other factors can be appreciated. The study of other branches of science should reveal why organisms have the forms they actually have—aerodynamics, for example, already provides explanations for many of the details of bird form. Elucidation of the ways in which growth is controlled so as to reach this form requires an experimental approach in conjunction with other biological sciences, particularly biochemistry and genetics.

Finally, it should be emphasized that, whatever the form of condensation, or statistical analysis, the summarized information, though adequate for the purpose of the original investigation, may not be an adequate substitute for the original data, if these data are subsequently required for other purposes. It is therefore important to take steps to preserve sets of measurements which are likely to be of general and continued interest. In cases in which the measurements are too extensive for publication in full, copies of the original data, or the original data themselves, should be filed so that they are available to other interested workers, as for example, in the archives established at the suggestion of the Committee on Biological Measurements of the British Association at the British Museum (Natural History) at South Kensington, or in the similar archives at the Royal Society of Edinburgh (British Association 1935). If the data have been analyzed by means of punched cards, a set of these cards should be preserved (new sets can easily be made from the original cards). At present there do not appear to be any central storage facilities available for this purpose.

THE ACCURACY OF GROWTH CURVES

By J. B. S. HALDANE, F.R.S., *Department of Biometry,*
University College, London

Most of the estimates of allometry in the literature are based on the regression of one variate on another. This is perhaps justifiable when body weight is taken as the independent variate, the weight of a small organ as the dependent one. It is certainly unjustified when, for example, the length and breadth of the same organism or organ are compared. The problem at issue is what line or curve

running through a scatter diagram should be taken as the best representative of the relation between the two variates. Kermack & Haldane (1950) point out that the line $\dfrac{x-\bar{x}}{\sigma_x}=\dfrac{y-\bar{y}}{\sigma_y}$ is invariant under a change of scale, and impartial between x and y. If this line does not pass through the origin, an allometric relation between x and y may be suspected, and if

$$X=\log x, \quad Y=\log y, \quad \text{then} \quad \frac{X-\bar{X}}{\sigma_X}=\frac{Y-\bar{Y}}{\sigma_Y}$$

may be taken as axis of allometry (Teissier 1948). If X and Y are normally correlated the coefficient of allometry

$$\alpha=\frac{\sigma_Y}{\sigma_X}=\sqrt{\left(\frac{\log(1+V_Y^2)}{\log(1+V_X^2)}\right)}=\frac{V_Y}{V_X},$$

nearly, where V_Y and V_X are the coefficients of variation of y and x. The corrections for non-normal correlation are worked out and shown to be small. The calculations are thus very simple, and it is believed that they are much more accurate than those based on regression.

If a number of individual growth curves are available, they can be averaged in two ways. The mean stature (or other measurement) at a given age can be plotted against age; or the stature (better the fraction of the adult stature) can be plotted against the age when this stature is attained. Neither method gives a representative growth curve. To obtain such a curve we should try to fit a small number of parameters to each growth curve. Thus if the rate of increase of the logarithm of human male height, or 'specific growth rate', is plotted against age, we obtain a curve descending to an asymptotic value which is small and negative, zero being reached at about 25 years. On this is superimposed a curve with a maximum about puberty and falling off rapidly on each side of it. Probably the whole curve could be adequately represented by four to six parameters, one of which would be approximately the age of puberty, another the total extra growth associated with puberty. A representative growth curve could then be obtained by using the mean, or better the median, values of these parameters. If several were found to be correlated, the number of independent parameters could be reduced. The parameters of such curves would have much more biological significance than those of curves representing averages. Polytokous animals in which litter size is an extra parameter, and animals in which weaning produces marked slowing, or even a temporary reversal, of growth are unsuited for such investigations. Men and cows would appear to be suitable. But it must be emphasized that a mean growth curve of either type, though it may be useful for some purposes, is a statistical artefact.

The mathematical analysis of form may be advanced by searching for highly correlated variates such as the lengths of bones in the hands, and replacing them by a general size parameter and a small number of shape parameters which may be ratios if growth is isometric, and allometric constants if it is allometric. Shape parameters, such as the ratio of lengths of adjoining metacarpals, are much less variable than, for example, the cranial index.

Some aspects of growth studies in fossils

By T. S. Westoll, D.Sc., *King's College, University of Durham*
(*Newcastle upon Tyne*)

Growth studies are frequently of use in solving problems of taxonomy in fossils. As a by-product of such a study on the Permian brachiopod *Dielasma elongata* (Schloth.), the occurrence of sexual dimorphism in brachiopods is thought probable. Various problems involving limited growth and seasonal (incremental) growth are discussed; some possibilities in the study of trilobite ontogeny are indicated. Ma's work on growth in corals has much significance in geological studies.

A review of recent work on relative growth in fossil vertebrates allows distinction between different types of allomorphosis to be drawn, and some of the methods are criticized. The dwarfing of various mammals is indicated as a valuable future study.

An interpretation of the evolution of *Gryphaea* is offered, which seems to afford an adaptive mechanism for a standard case of recapitulation and orthogenesis. Certain 'katagenic' trends in Graptoloidea and ammonites are interpreted as paedomorphic developments in forms of increasing potential size; the idea of a *potential hypermorph* is used.

Finally, an analysis of growth in the skull of a labyrinthodont is used as a basis for discussion of changes in proportion in the evolution of major groups of vertebrates.

Introduction

Fossil remains frequently set problems that involve investigating some aspect of the growth of individuals. This means generally the question of changes with change of *size*, since in most cases the duration of the individual development to a particular size cannot be determined. In some cases 'annual' or other cyclic growth can be demonstrated, but it should be remembered that the duration of a Cambrian 'year', for example, may perhaps have differed somewhat from a present year, though probably not very significantly.

There are great difficulties in the widespread use of accurate growth studies among fossils—large collections may not be available, and may often come together in museums from somewhat different localities and from a certain range of strata; even where numbers of specimens from one locality and a single stratum are known, they form often a selective sample of the life history or of the population at that time, representing only a fraction of the individuals living in a particular range of habitats; and, of course, only hard parts of the skeleton are normally preserved, and these may show much distortion, destruction, and mineral replacement, while removal of the matrix may be very difficult.

Earlier palaeontologists used growth studies rather extensively in the attempt to set out the phylogeny of various groups on the basis of the 'biogenetic law', and to this day a great part of the text-book classifications of many groups is modelled on the results of such work. Many palaeontologists have been far less inclined than zoologists working on living creatures to discard the recapitulation principle, and one reason for this is very pertinent to this essay. Embryologists, who have most violently opposed the general validity of the 'biogenetic law', have, in general, studied *far earlier* growth stages than those available in by far the majority of fossils. Changes from 'juvenile' to 'adult' stages are often not fully recorded for living organisms, and it is just this range that interests the palaeontologist.

Some palaeontologists, however, notably Spath (1933, 1938), have also produced intensive studies discrediting the 'law'. It remains true that the use of growth stages, *combined with sound stratigraphy*, is essential for phylogenetic studies in many groups (e.g. corals, ammonoids, lamellibranchs, brachiopods and trilobites). There are now very few palaeontologists who would draw far-reaching phylogenetic conclusions from ontogenetic studies alone.

Cases of taxonomic and biological significance

Collections of fossil materials from restricted horizons have frequently been investigated by techniques familiar to zoologists. Single cross-sections in time of a fossil lineage do not tell us more, fundamentally, than study of living species can do. They are nevertheless useful in taxonomic and stratigraphical studies; what were formerly recognized as distinct species, even placed in separate genera and (in some cases) families, have been determined subsequently as growth stages of the same species. Whether the growth question is involved or not, it is now generally recognized that statistical study of assemblages of fossils is essential for critical work. The selective nature of preservation and collection often makes growth studies impossible, but an example from the writer's own experience shows the taxonomic importance of, and the occasional emergence of an unexpected problem arising from, simple considerations of size increase.

The small brachiopod *Dielasma elongata* (Schloth.) and its 'varieties', among which more than one species was formerly recognized, is abundant in the Permian reef-dolomites of Co. Durham. Analysis of variation of several measurable characters showed that, according to the evidence of scatter plots and frequency of occurrence of ratios of breadth/length, Bd/L, Td/L,* etc., only one species is present. But the thickness/length ratio showed a peculiar anomaly. Below a certain size a typical unimodal distribution is found; above that size the distribution becomes at first very asymmetrical and then bimodal, so that a 'thick-form' and a 'thin-form' group can be recognized; the latter persists to a greater total length. Growth stages of the individual shell are preserved as growth lines; corresponding prominent growth lines can be traced on each valve, so that a reconstruction of the individual at a few growth stages can be made. It is found that the 'thin-form' group remains relatively thin throughout life, but the 'thick-form' shells, which are initially 'thin-form', tend to grow more rapidly in thickness above a size of *c.* 6 to 8 mm. total length. The 'thin-form' and 'thick-form' shells $> c.$ 8 mm. length are about equal in number, and this may be regarded as evidence suggesting sexual dimorphism, which, so far as the writer is aware, has not previously been noticed in the group. Some other brachiopods which have been statistically surveyed in an appropriate manner (*Terebratula punctata*, Alkins, 1923; *Reticularia lineata*, Day 1915, Alkins 1920; *Rhynchonella boueti*, Alkins 1923, Aitken & McKerrow 1948) show little or no sign of such a thickness dimorphism; but others show at least some tendency to a platykurtic, asymmetrical or peak-and-shoulder distribution of thickness/length or thickness/breadth ratios. In 600 *Pugnax* from the Car-

---

* Bd, Td = distance of maximum breadth (Bd) or thickness (Td) from hinge-line; L = total length.

boniferous Limestone of Castleton, Derbyshire, B. Simpson (1933) found, as in *Dielasma elongata*, no evidence of heterogeneity in the length (height)/breadth ratios, but again a 'thick-form' and a 'thin-form' group were distinguished, in this case as two species (*Pugnax pugnus* Mart., and *Pugnax* cf. *sulcatus* (J. de C. Sow.)). It seems highly probable that sexual dimorphism would offer a likely alternative.

Some special problems are posed by animals showing a limit to growth. Thus in fossil mammals measurements of tooth crowns and of fully ossified parts of skeletons
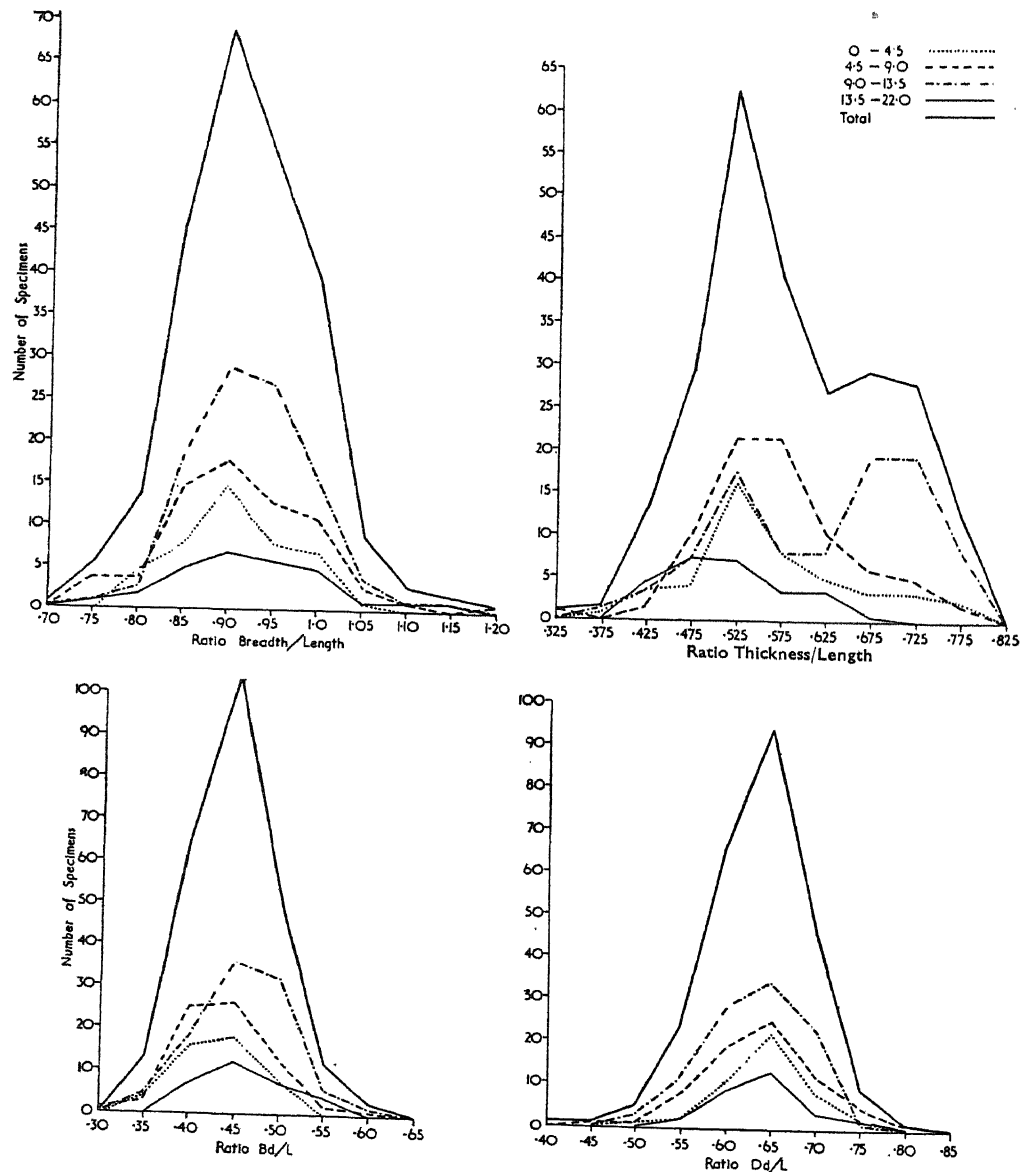
FIGURE 22. Frequency polygons of four ratios measured in a collection of *Dielasma elongata* (Schloth.). *Bd, Dd* = distance of maximum breadth and thickness respectively from posterior end of shell; *L* = total length of shell. The data are analyzed into size groups based on total length.

are widely used in taxonomic studies, for example, in comparing samples of faunas of different ages or from different localities. It has also recently been shown that certain early vertebrates (ostracoderms) apparently ceased growing in size at a certain stage, and that dermal ossification began then. Prior to this period of ossification the only hard skeleton was probably in the form of skin-denticles, possibly as composite associations in some forms (Westoll 1945; Denison 1947). This seems to have far-reaching taxonomic significance, since the whole status of the Coelolepida becomes dubious; it also brings out a series of new problems, of a physiological and ecological nature, that may never be susceptible to solution.

Seasonal and incremental growth

Incremental and seasonal growth is important in many palaeontological problems. Since the time of Barrande it has been known that trilobites show a familiar arthropod pattern of growth and ecdysis, and in recent years Raw (1925), Stubblefield (1926) and Størmer (1942) in particular have contributed much to our knowledge. An indication of the way in which growth measurements may be applied during an investigation is provided by further analysis of the data on *Leptoplastus salteri* (Call.) provided by Raw (1925); the post-protaspid stages were then described, with a useful table of measurements. During these growth stages the total number of post-cephalic segments increases from 7 to 15, and the number of free thoracic segments (the 'degree') from 0 to 12. It is likely that during this part of the growth cycle the relative thickness of the body did not vary very disproportionately, so that (length × breadth)$^{\frac{1}{2}}$ probably gives a good measure of relative volume at each stage. Raw gives measurements of L (total length), L_C (length of cranidium or cephalon), B_T (max. breadth of thorax), and B_C (max. breadth of cranidium), from which B_A (arithmetic mean of B_T and B_C) can be determined. The volume relations $(LB_A)^{\frac{1}{2}}$ and $(LB_C)^{\frac{1}{2}}$, and for the cephalon $(L_C B_C)^{\frac{1}{2}}$, have been calculated for each 'degree'. The holaspid range, after the attainment of the full number of free thoracic segments, clearly includes several stages; the same may perhaps be true of each degree, and there may be several stages missing from the collection. These possibilities may now be explored. It is also necessary to decide whether, as a first approximation, each instar or 'degree' was represented by individuals of the same size.

In figure 23 A, the successive 'degrees' of ontogeny are plotted against the logs of $(LB_A)^{\frac{1}{2}}$ and $(L_C B_C)^{\frac{1}{2}}$. No example of degree 8 was known to Raw. It can be seen that the points lie very close to straight lines, those representing degrees 1 to 4 inclusive on lines of half the gradient of 4 to 7 and 9 to 12; that the volume of degree 7 is almost exactly the same as that of degree 9, so that degree 8, if it actually existed, was probably very similar in size; and that the distribution of the points suggests very strongly that at any particular 'degree' up to 12 the individuals were probably much the same size, otherwise an array of points of this sort would represent a highly improbable sampling error. The relationship between the gradient of the lines allows a different plotting to be made so that all the points lie close to straight lines of the same gradient, with a discontinuity from degree 7 to degree 9 (figure 23 B). This could be interpreted as due to the absence from the

collection of certain instar stages, with a nearly constant volume increase between each two instars, except in the critical 7 to 9 stage, which could well be an expression of the changes due to the onset of sexual maturity. This point could probably be decided by collection of extra material; if no additional stages up to degree 12 were found in a reasonably large second collection, the probability that the first plot gives a good record would be very high. It would also perhaps allow one to decide whether degree 8 had any real existence, or whether degree 7 produced, after ecdysis, *two* additional free segments in the thorax. On the close plotting (figure 23 A) the early curve $(LB_C)^{\frac{3}{2}}$ (degree 1–4) represents a volume increase of 1·45:1 per ecdysis, while
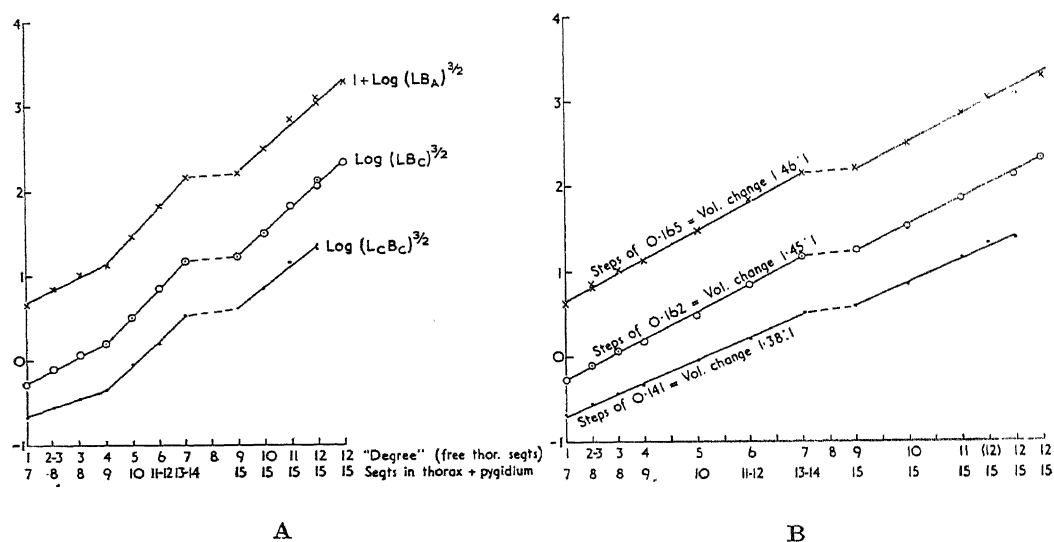
FIGURE 23. Relationship between 'degree' and a function of volume in the trilobite *Leptoplastus salteri* (Call.). L = total length, L_C = length of cranidium, B_A = arithmetic mean of breadth of cephalon and thorax, B_C = breadth of cephalon. Data from Raw (1925). In A, 'close-plot', in B, 'open-plot'.

the degree 4–7 and degree 9–12 ranges represent a volume increase of about 2·1:1 at each ecdysis, which recalls Przibram's 'law' of doubling of cell number at each ecdysis. The cephalon shows a volume increase of 1·38:1 and 1·9:1 in the corresponding ranges. The more 'open' plotting gives ratios of 1·45 or 1·46:1 for the whole body and 1·38:1 for the cephalon throughout, and it may be significant that the lobsters recorded in the British Museum *Guide to the Crustacea* (1927, p. 18) show a relation between (length)[3] at various moults indicating a volume change per moult of roughly 1·4:1. On this basis perhaps 17 to 18 ecdyses might be required to produce a small degree-12 individual from the earliest known meraspid stage, with perhaps 20 to 25 'adult' ecdyses to produce the largest known individuals; the closer plot would allow 11 and 10 to 12 respectively to be considered likely. It is finally necessary to note that the poorer 'fit' of larger individuals could well be due to progressive 'spread' of variation in size at corresponding instars, but could easily be due also to fundamental error in this means of approach; once more, larger collections would probably allow this point to be settled.

Comparison of these results with those of Stubblefield (1926) on *Shumardia* and Whittington (1940) on *Onnia* is valuable, even though the former does not give full tables of measurements. In neither case is there an increase in number of segments without notable increase in size, just before the formation of the final total number of segments, so far as can be seen from the data. Both genera show that there are probably two (or possibly more) ecdyses within each 'degree', and preliminary plotting of the data in the same manner as for *Leptoplastus* shows that the points can be arranged to lie close to straight lines over a considerable part of the range with slopes indicating a volume increment of *c.* 1·4:1 at each 'step'. This is close enough to the figures for the 'open' plotting in *Leptoplastus* to warrant further examination, but the writer is only too fully aware of the subjective nature of the plotting and these results may well be illusory. Examination of the material should settle this point, and may throw further light on other growth problems which need not be mentioned here.

A most ingenious and original use of seasonal growth has been made by Ting Ying H. Ma (1934–43) in his studies of fossil and recent corals. He found that reef-building corals in the seas south of Japan frequently show a very marked seasonal growth of the skeleton. Where the average water temperature is low, the annual growth of the skeleton is small, but very clearly distinct 'summer' and 'winter' additions may be seen. In warmer conditions the skeleton is less compact, and the broader seasonal bands may be less obvious. In the equatorial region, where seasonal variations of temperature are little marked, growth is probably most rapid but shows little or no seasonal variation. Since corals are plentiful only in relatively warm shallow seas, these factors help to mark out an 'equatorial' and two flanking belts, outside of which corals will not be numerous. In raised-beach deposits of the Japanese area it was found that the coral fauna, composed almost entirely of living species, showed growth patterns of a different 'zone' from that now characteristic of the neighbouring seas, which presumably reflects changing Quaternary climates. Applying the same methods to fossil corals of certain earlier geological periods, Ma finds that it is possible to indicate rather closely the position of the Ordovician, Silurian and Devonian equator in certain parts of the world. These do not coincide with the present equator, and in the Devonian period, for example, the equator so determined crosses both North America and Eurasia; the tracks on the present continents do not lie on the same great circle, but come into proper relationship on what is essentially the Wegener-Du Toit displacement. These results are so promising that wider work in the same field is urgently needed.

An interesting phenomenon in early Crossopterygii and Dipnoi has been interpreted as evidence of incremental growth. In many individuals of, e.g. *Osteolepis* and *Dipterus*, the superficial layer of cosmine on the dermal bones and scales is very extensive, and may completely obliterate sutures. In such individuals it is difficult to see how growth could have taken place without extensive changes in the cosmine, and yet quite small individuals (e.g. in *Megalichthys*) may show this continuous cosmine layer. Westoll (1936) showed that the cosmine could be resorbed, and that when this histologically distinctive skeletal tissue was lost the

fishes concerned looked so different that they had been placed in different genera and families. The more recent works of Bystrow (1942) and Jarvik (1948, 1950) have corrected some important errors in the original interpretation, and make it possible that the process of resorption and reformation of cosmine was, in Dipnoi, not necessarily periodic and possibly seasonal, but in some cases was more or less continuous. On the other hand, Jarvik's work (1950) suggests that in Osteolepids the process was rapid, taking place at intervals. The writer once thought that the cosmine was used as a portable reserve of calcium phosphate during rapid growth; it is perhaps more probable that the resorption was a physiological device (a sloughing process) for ridding the body of excessive calcium salts, and that when more adequate physiological control of calcium intake and deposition was evolved (as is indicated also by extensive reduction of general ossification) there was no further need for cosmine to be formed and destroyed.

Allometric studies in fossil groups

The main contribution of growth studies on fossil material involves the occurrence of time sequences, based on stratigraphical succession. An important generalization is that many evolutionary sequences, in invertebrates (Newell 1949) as well as vertebrates, show progressive increase in maximum size. Newell gives good examples of allometric studies in invertebrates. Even without any marked dysharmonic allometric size changes, there is probably an effective upper limit of size for any group, controlled by mechanical, physiological or ecological factors. But many large animals show 'hypertrophy' of some features (e.g. horns, tusks, etc.) which follow fairly closely the simple allometric formula of relative increase ($y = bx^{\alpha}$, often written also $y = bx^{k}$), and further growth on the same law would in many cases lead to unviable types. Furthermore, as Hersh (1934) has noted for mammals, the size at birth, which in general increases with maximum adult size, might become so large as to cause difficult parturition if such structures were already well developed.

Many changes previously quoted as evidence of orthogenesis have been shown to be capable of analysis as effects of general size increase, e.g. the horn development of Titanotheres (Osborn 1929; discussed by Hersh 1934, etc.).

Comparison of a succession of phyletically linked fossil forms with growth stages of one of the members (i.e. allomorphosis v. heterauxesis, Huxley, Needham & Lerner 1941) may be very illuminating. If ontogenetic recapitulation were perfect and the lineage closely defined, the relative growth curves for a late member of the series would coincide with the corresponding curves for the adult members of the lineage. So far as the writer is aware, no case of this kind has been discovered; it would be surprising if it were. However, some 'near approaches' have been suggested. Robb (1935) graphed the ratio length of face : length of skull in (a) ontogenetic stages in the living horse, Equus caballus, (b) adult skulls of different races of horses, and (c) fossil horses from Hyracotherium (Eohippus) to Equus. All points fall within a very narrow band, but the linear regressions for each set, on a log/log plotting, do not exactly coincide. It is, of course, inexact to have one of the variables included in the measurement of the other, since this

minimizes changes; Reeve & Murray (1942) have reinvestigated this case, and find that (plotting log face length (y) against log cranium length (x)) the *Equus* ontogeny shows a decided change, with $\alpha = 1 \cdot 5$ up to a cranial length of 15 cm., and $\alpha = 1 \cdot 0$ above that size. The phylogenetic line from *Hyracotherium* to *Merychippus* gives a gradient $\alpha = 1 \cdot 8$, and the face is proportionately shorter in this line than in young *Equus* of corresponding cranium length. *Pliohippus* and some species of *Merychippus* lie on the *Equus* ontogeny line; 'aberrant' hypsodont horses such as *Plesippus* lie off the *Equus* line. There is good evidence that the 'main line' to *Equus* shows a decided break at about Miocene times, i.e. when the cheek teeth became hypsodont; the phase of molarization of the premolars, in the late Eocene, seems to have little effect on relative growth of the face. *Hypohippus*, the later Miocene-Pliocene 'forest horse', which retained low-crowned teeth, has a relatively short face. So much has been written about the evolution of the horse, and so many erroneous ideas persist in text-books, that a large-scale project to study relative growth in living and fossil horses, on as large an array of characters as possible, is very desirable. There is probably sufficient material in various American museums to allow the terminal part of the ontogenetic curves for several extinct forms to be drawn. It is already clear that horses did not run their evolutionary race in orthogenetic blinkers (e.g. Stirton 1940; Camp & Smith 1942; Simpson 1944; Romer 1949). It is now necessary that we should know more about the critical changes in each lineage. The excellent discussion of living anteaters by Reeve (1940, 1941) would be of great methodological value in this work.

Other studies on relative growth in fossil vertebrates (e.g. by Hersh 1934, on titanotheres; by Phleger 1940, on felids and merycoidodonts; by Phleger & Putnam 1942, on *Merycoidodon*; by Gray 1946, and Lull & Gray 1949, on Ceratopsia) have provided valuable information, and made it clear that growth studies can be of great help in some phylogenetic problems. Some general conclusions have been drawn that require further comment. Several different types of relative growth curves have frequently been confused:

(*a*) the ontogenetic relative-growth (heterauxesis) and

(*b*) absolute-size allometry (allomorphosis) of essentially adult material, of several different kinds, e.g.

(i) members of different contemporaneous stocks or races of a species (race allomorphosis),

(ii) members of different species of the same genus, without paying attention to differences (usually small) in geological age (species-form allomorphosis, and similarly genera-form allomorphosis, etc.),

(iii) members of a lineage, showing progressive change in size, and of known relative geological age (lineage allomorphosis).

Distinctive names for the three main types of allomorphosis are here suggested.

In fossil collections it is often impossible to decide whether material of different size represents terminal samples of narrowly similar ontogenies, or individuals of different races or genetic stocks. The measurements of *Smilodon californicus* and *Felis atrox* (Phleger 1940) are plotted in a broad band, suggesting that the parameter *b* differs in different individuals of about the same size. Some remarkable

aspects of race allomorphosis contrasted with genus-form allomorphosis in brain-size relationships are discussed by de Beer (1940a). Lumer (1940) has shown that analysis of race allomorphosis in different breeds of dogs allows six allometric tribes to be distinguished; it would be desirable to compare these results with heterauxesis studies.

There is always the additional difficulty of sex-linked differences; there is some discussion of this point in Gray's (1946) account of *Protoceratops*.

Hersh (1934) suggests that each genus of Titanotheres is characterized by its own parameters *b* and *k* (α) (this is essentially species-form allomorphosis). Phleger & Putnam (1942) show that 155 specimens of *Merycoidodon* allow straight regression lines to be drawn for six pairs of characters in log/log plots (species-form allomorphosis), while each of the four species represented by more than a dozen specimens (the remaining species had a single representative) has its own regression line (? terminal heterauxesis combined with race allomorphosis), differing from the 'generic' line. It may be noted that some of Hersh's paired characters are wrongly chosen, since the greater measurement includes the less (e.g. upper molars/upper cheek teeth), and that most of the longitudinal measurements used by Phleger (1940) in his analysis of felids and merycoidodonts are plotted against maximum length of skull. It may be suspected that most of the cases of 'generic' allometry curves (from individuals showing species-form allomorphosis + race allomorphosis + terminal heterauxesis) involve species not in a single line of descent, and the basis for such regression curves is fundamentally unsound.

Gray (1946) has studied relative growth in growth stages of the early ceratopsian dinosaur *Protoceratops*, and has also drawn what are approximately lineage-allomorphosis curves for several corresponding pairs of measurements on the skulls of members of the 'long-crested' and 'short-crested' Ceratopsia. Lull & Gray (1949) applied to the same material the d'Arcy Thompson method of a deformed Cartesian grid. They point out that such deformation summates graphically the relative growth changes in many separate measurements, and that each line segment may follow the simple allometry formula, $y = bx^{\alpha}$; and also that progressive deformation (on what appears to be a phyletic trend) often cannot be extrapolated very far without the grid coming to a critical condition (e.g. one side of a triangle becomes equal to the other two sides). This suggests that there may be real limits to an 'orthogenetic' trend involving changes of proportion. Lull & Gray also distinguish between growth curves for a phylogenetic 'adult' series (lineage allomorphosis as defined above) and ontogenetic growth series (heterauxesis).

In spite of the criticisms that have been made of these studies, there can be no doubt that they have use in phylogenetic problems. The programme of research on fossil and recent horses, outlined above, would undoubtedly prove even more valuable as a basis for further work.

Of particular interest is the work of Romer & Price (1940) and Romer (1948) on fossil Pelycosauria. Seeking some unit of measurement for these animals, they concluded that (total weight)$^{\frac{1}{3}}$ or (total volume)$^{\frac{1}{3}}$ would be suitable, and that the cross-sectional area of the trunk vertebral centra (which form the main weight-

supporting structural member in the trunk) would be proportional to the total weight. Thus $r^{\frac{2}{3}}$, where r = transverse semi-diameter of a thoracic centrum, should be a useful measure of size,* and is termed the *orthometric linear unit* (o.l.u.).

This unit proved invaluable in a taxonomic revision of the pelycosaurs (Romer & Price 1940), and has recently (Romer 1948) been applied to relative growth problems. Thus in species of *Ophiacodon*, the length of longest neural spines (y) plotted against the orthometric linear unit (x) shows linear relations, $y = c. 10x$. But in approximately phyletic sequences of species of long-spined genera of Sphenacodontia the same variables show striking differences:

	b	α
1. *Sphenacodon ferox*→*S. ferocior*→*Ctenospondylus*	0·688	3·22
2. *Dimetrodon*, a sequence of species	39·8	1·6

The other long-spined pelycosaurs, the Edaphosauria, are also represented by an approximately phyletic series, for which the corresponding regression is $y = 15\cdot34\ x^{0\cdot58}$; the apparently negative allometry of *length* of spine in this case is in some ways misleading, since the spines show the progressive development of lateral processes which increases their bulk and surface area. The enlarged spines were clearly embedded in a thin 'sail' of soft tissues supplied with blood vessels; Romer points out that the areal development of the sail in *Dimetrodon* is remarkably close to what would be expected if the 'sail' functioned as a thermo-regulating device. In *Edaphosaurus* the sail had apparently already developed to nearly its maximum possible relative size (mechanically) when the first well-known species of the genus were alive. Romer also shows that the disproportionately enlarging skulls (as measured in o.l.u.) of *Ophiacodon* ($\alpha = 2\cdot3$) and *Dimetrodon* ($\alpha = c.\ 2$) can be accounted for on functional grounds.

The general increase in size shown by many of these phyletic lines is a well-known phenomenon, which has received many 'explanations' which at best are only very partial ones: e.g. Haldane's (1932) suggestion of the selective advantage in pre-natal growth in mammals which produce many young at a birth; the effects of selection of large males in species where the males fight for possession of the females, and so on. Recently D. M. S. Watson (1949) has shown, in brief, that a larger animal on the whole is able to make more efficient use of its food, but that it needs to eat more and oftener to keep alive and active. Various changes in the dentition, besides others in the digestive tract not open to palaeontological analysis, can be seen as necessary consequences of increase in size, and are often understandable on quite simple applications of the power relations of growth in length, surface area and volume. A corollary is that there is probably always a limit to the effective maximum size of a group.

It may be noted here that Hersh (1934—free length of nasals in titanotheres against skull length) and Gray (1946—length of nasal horn core in later Ceratopsia against skull length) have noted cases in which certain measurements actually decrease with increase of body size, and Hersh calls this hyperbolic heterogony; Reeve & Huxley (1945) call it enantiometry. Hersh has also endeavoured to find

* Strictly speaking, $cr^{\frac{2}{3}}$, where c is the cube root of the unit of linear measurement, e.g. 1 mm.$^{\frac{1}{3}}$.

a relationship between b and α in allomorphosis curves (?species-form) for similar variables in genera of titanotheres, having noted that, as is very general, b decreases as α increases. His suggested formula is $b = Be^{-r\kappa}$, where B and r are constants for each set of corresponding allomorphosis curves.* Lumer (1940) has shown that such a relationship may be an artefact of the method, but it often holds.

Dwarfing, while less common than size increase as an evolutionary phenomenon, is known in many lines, and is a neglected field of study in growth changes. What is probably 'environmental' dwarfing is often found in fossil invertebrates in certain kinds of rocks, and is usually thought to reflect unfavourable conditions. Thus the fossils of the upper part of the Permian reef dolomites of Durham show striking diminution in variety and size. There are some amusing consequences. *Dielasma elongata*, for example, is much smaller than normal, and its brachial loop does not develop beyond the centronellid stage, so that a perfect reversal of the 'biogenetic law' is found from base to top of the reef (cf. Trechmann 1925).

Much more striking examples of dwarfing of a different kind are to be found in mammals. Thus there are several dwarf horses more or less closely resembling larger contemporaries; the middle Miocene *Archaeohippus*, for example, is essentially a dwarfed *Parahippus*; *Nannippus* (Lower Pliocene) essentially a dwarfed *Hipparion*. Though measurements have not been analyzed, these dwarf forms are more nearly *miniatures* of larger forms than arrested stages in normal ontogenetic growth (heterauxesis). Much the same is true of the fossil pygmy elephants of the Mediterranean islands; the small *Elephas falconeri* from Malta, only 3 ft. high at the shoulder, is a miniature *Palaeoloxodon*, and our limited information suggests that these animals had tusks larger than growth stages of 'normal' *Elephas* of the same general size. There is, at any rate, complete certainty that such dwarfing is not a size reversal on the lineage-allomorphosis track; all these dwarfs are far more 'advanced' in structure than ancestral forms of similar size on the main line of descent.

Ontogeny and phylogeny

De Beer (1940 b) has given an admirable account of the relations of ontogeny and phylogeny; his figure 2 shows at a glance the various phenomena: *retardation*; *adult variation, neoteny* and *paedogenesis* (together classed as *paedomorphosis*); *hypermorphosis*; and *acceleration*. Many of these have bearings on palaeontological problems, and de Beer quotes examples from graptolites, *Gryphaea* and ammonites.

For example, many groups showing phyletic increase in size provide examples of recapitulation; this is stated in other terms in de Beer's definition of hypermorphosis. The writer has found the concept of a *potential hypermorph*† to be useful. In such lineages as those just mentioned, the potential hypermorph at any stage is the result to be expected on moderate increase in size beyond the normal maximum, on the laws of growth characteristic of the individual; it would, in general, resemble immediately following members of the lineage, but in detail

---

* Hersh gives r a negative sign in his worked examples, which is incorrect; the negative sign already appears in the formula.

† There should be no confusion between this term and the term *hypomorph* used by some geneticists to describe genes with decreased effect compared with wild-type alleles.

there would be adjustments to environmental changes, to new genetic and other possibilities, etc., such as probably always occur in the slow process of evolution.

Modifications of this kind help to explain why a lineage allomorphosis is not strictly parallel to any of the numerous ontogenetic heterauxeses that produce it. It is, however, in cases of paedomorphosis that the concept of the potential hyper-morph is most useful.
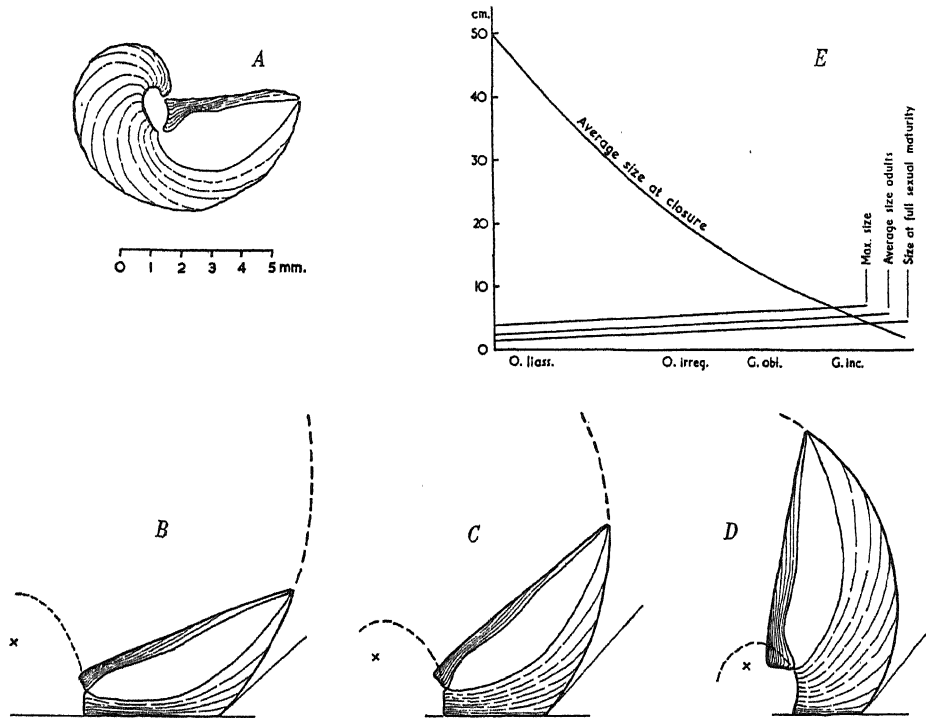
FIGURE 24. *A*, section through specimen of *Gryphaea incurva* showing thickness of shell and arrangement of growth lamellae, semi-diagrammatic, from specimens in Department of Geology, King's College, Newcastle upon Tyne. *B*, *C*, *D*, diagrams showing effect of shortening of area of attachment, and slight increase in angle of rise of profile (40° in *B*, 45° in *C*, 50° in *D*) and spiral angle of external curve (60° in *B*, 65° in *C*, 70° in *D*) on shell shape and thickness in a *Gryphaea* lineage, the dimensions of the body-cavity being the same in each sketch. *E*, time-changes in average size at 'closure', maximum size and average adult size of individuals, and size at full sexual maturity, in the *Gryphaea incurva* evolutionary development; geologically older assemblages to left, latest to right.

One of the most famous fossil records of evolution is that of the oyster *Gryphaea*. Elementary statistical analysis of collections made from successive 'banks' of the oysters in the Lower Lias allowed Trueman (1922, 1930, 1939) to make far-reaching conclusions about the taxonomic significance of progressive shift of the mode in a variable stock traced through a thickness of strata. Swinnerton (1938, 1939, 1940) has given much attention to the evolution of these modified oysters in a series of thought-provoking presidential addresses. The 'gryphaeoid trend' has operated several times during the Mesozoic era (Arkell 1933). The writer proposes here to discuss what is clearly the crucial point in the evolutionary history of the *G. incurva* group—the incurving of the left umbo to such a degree that the right

valve, in the most advanced stages of the trend, can have had only a very limited range of opening movement.

The group is regarded by Trueman (1922) as derived from *Ostraea liassica* Strick. through *O. irregularis*; Swinnerton (1940) recognizes two 'lineages', one *O. irregularis* → (*Gryphaea dumortieri*) → *G. obliquata* → *G. incurva* var. *crassirugata* → *G. incurva*, and the other *Ostraea irregularis* → *Gryphaea* cf. *obliqua* → *Gryphaea* cf. *incurva*, *G. macculloughi* and *Gryphaea* cf. *cymbium*. *Ostraea liassica* is a flat oyster, attached by most or all of the surface of the left valve, and normally 2 to 3 cm. in length (max. *c.* 4 cm.). In succeeding zones the 'average' oyster shows a progressive reduction in the area of attachment and a greater degree of coiling of the left valve, and also increase in size, so that in *Gryphaea incurva* (large individuals up to 6 cm. or even more in length) the modal length of the attachment area is *c.* 1·5 mm. (occasional individuals show lengths up to 25 mm.) and the modal number of whorls in the 'coil' is 1 to 1¼ (data from Trueman 1922, and Swinnerton 1939). During this process the valves (especially the left valve) become very thick.

The coiling is usually in profile, a fairly close approach to a logarithmic spiral, as first noted by d'Arcy Thompson; the spiral angle is lowest in the earlier members, increasing to about 80° in *G. incurva* (cf. Trueman 1922; Maclennan & Trueman 1942); Trueman also notes that the 'effective' spiral angle normally increases during ontogeny. The 'angle of rise' or 'slope of area' between the area of attachment and the first free part of the left valve, used by Swinnerton in his comparisons, is normally less than the spiral angle, since the centre of the logarithmic spiral is not in the plane of the area of attachment. Swinnerton's figures show that the angle of rise of profile shows great range of variation, but the mode is about 40 to 50° in fairly early and quite advanced forms.

These changes have now become a standard example of 'orthogenesis' of a type often 'explained' by recourse to some mysterious predestination or other esoteric cause. de Beer (1940*b*, pp. 71–73) classifies it as a case of hypermorphosis combined with acceleration. A possible functional explanation, involving natural selection and based on study of the growth processes, may be outlined here.

The large size of the attachment area in early members indicates that animals spent at least most of their lives fixed to rock surfaces, shells, etc. The later members clearly broke away from such surfaces at an early stage of growth, and probably lay in more or less crowded assemblages on and in the yielding sediments of the sea floor. The thickened shell is advantageous in such circumstances, holding the shell more or less in position during the action of currents. In *Ostraea liassica* and *O. irregularis* the shell is not very greatly thickened. During growth, new shell material was laid down on the inner surface of each valve, and the shell was prolonged by the secretion activity of the edges of the mantle. When the first 'free' part of the left valve develops, the ratio between the rate of thickening of the ventral part of the valve, and the rate of advance of the ventral margin, determines the 'angle of rise of profile', which shows considerable variation, with a predominant range of say 35 to 50°. The early growth stages of the left valves of later forms (*Gryphaea obliqua*, *G. incurva*) very closely resemble in shape the

normal fully grown left valves of *Ostraea irregularis*, as Swinnerton (1939, 1940) has very clearly shown; thus *Gryphaea incurva* at roughly one-quarter the full length (measured radially from the left umbo) is essentially a diminished replica of *Ostraea irregularis*, just as *O. irregularis* at one-third to one-half adult length resembles *O. liassica*. Put in another way, the majority of specimens of some of these assemblages have lengths of area of attachment less than the following figures:

O. liassica	say 25 mm.	*G. obliquata*	say 5 mm.
O. irregularis	say 10 mm.	*G. incurva*	say 2·5 mm.

The resemblances between appropriate growth stages are indeed close, as Swinnerton has shown. We may therefore suppose that if *O. irregularis* or *O. liassica* had been able to grow very large on the law of growth already determined, they would have attained a *Gryphaea incurva*-like form at about 4 and 10 times (respectively) their normal adult length, i.e. at say 10 to 12 and 25 cm. respectively. During this growth the shells would have become extremely thick.

Though Swinnerton's graphs (1939, figures 4*b*, 5*b*) of the distribution of 'angle of rise of profile' show modes at 40° in both *Ostraea irregularis* and *Gryphaea incurva*, the skewness of the latter is stronger, and the arithmetical means calculated from his curves are about 40° and about 50° respectively. This means that in *G. incurva* the shell at the *Ostraea* phase of growth is probably thickening slightly more rapidly with growth in length, on the average, than the adult of *O. irregularis*. This implies that the spiral angle of the 'potential hypermorphic' gryphaeoid development of *O. liassica* would be less than that of *O. irregularis* and still less than that of *Gryphaea incurva*, which in turn implies that the *G. incurva*-like stage would be reached at still larger sizes than the *c.* 10 to 12 cm. (*Ostraea irregularis*) and *c.* 25 cm. (*O. liassica*) deduced above. Now *Gryphaea incurva* populations could not have grown very much larger in size because of the imminence of closure of the valves. We may suppose that at lengths of say 50 cm. (*Ostraea liassica*), 20 cm. (*O. irregularis*), 10 to 15 cm. (*Gryphaea obliquata*), and 6 to 8 cm. (*G. incurva*) shells growing on these laws would, in the greater part of the population, fail to open.

It is here suggested that *Gryphaea* was highly adapted for life on unconsolidated sea-floors, and that early thickening of the shell was an important factor in this adaptation. Spat, settling on some shell or hard surface, grew until the shell was fairly thick and massive and broke free; the greater the thickening, the smaller, on the average, the area of attachment. The result of this is the tightening of the spiral by (*a*) shortening of the area of attachment and (*b*) the increase in the effective spiral angle. It is significant in this respect that Trueman (1922) noted an ontogenetic increase in the spiral angle in *Gryphaea*; the requirements of young individuals, still attached, are different from those of free-living older individuals.

While this process of tightening the spiral and decreasing the closure size is going on, the 'adult' (and maximum) size of the individuals in the populations is rising from say 2·5 (3 to 4) cm. in *Ostraea liassica* to 5 to 6 (6 to 7) cm. in *Gryphaea incurva*. In *G. incurva*, as already noted, very large individuals cannot have been

able to open their valves very far, and were probably inefficient organisms. Loss of the largest members and heaviest feeders would have no deleterious effect on the population as a whole—rather the reverse if competition for food in the oyster-beds were intense. So there would be no selection *against* this process until it began to affect forms not yet sexually mature, or not yet functionally female (if *Gryphaea* was a protandric hermaphrodite like many oysters). This might gravely affect the survival chances of the group.

It is interesting to note that occasional *G. incurva* of large size have a 'flared' ventral part of the longitudinal profile (Trueman 1922, figure 6; Maclennan & Trueman 1942, figure 3), indicating a sharp drop in the effective spiral angle, and a marked reduction in the rate of thickening with increasing size. This is probably a significant point suggesting some functional difficulty in opening the shell; if any genes controlling such an effect existed they do not seem to have become established in the populations.

It would seem that a theory on these lines affords a mechanism, controlled by natural selection, not only for an orthogenesis ending in extinction, but also for one of the most carefully described and fully discussed examples of recapitulation in fossils.

The Graptoloidea, as an example of a group of colonial organisms, afford interesting evolutionary histories.

Early Graptoloidea have numerous stipes, each produced by repeated budding of thecal individuals from the embryonic sicula or from other thecal individuals.* Even in the early graptoloids two important facts may be noted—there is frequently a tendency for the thecae along each stipe to increase somewhat in size from the most proximal theca, but there is also apparently a maximum possible thecal size for each species (or race—geographical races may show different maximum sizes). Graptoloids present many interesting problems that can be approached from the consideration of growth stages; attention will be confined here to some problems, affecting the latest and structurally simplest forms, the Monograptidae, in which the potential hypermorph concept is valuable.

The 'biform' monograptids (cf. Bulman 1933) have the proximal thecae more or less strikingly distinct in nature from the more distal ones. In the members of the 'hooked' trend, the early ('anagenic') members are biform with hooking of proximal thecae, the most typical and fully developed members (e.g. *Monograptus priodon*) are not biform and all thecae are hooked, and the later ('katagenic') members are again biform, the distal thecae gradually losing the hook. The whole series may be regarded as potentially hooked if the thecae could grow big enough (figure 25), but there is usually a well-marked limiting size of actual thecae. The 'katagenic' members indicate that the fully hooked potential hypermorph would have been progressively larger. In the 'isolate' line, also fully discussed by Bulman, the same concept can easily be applied, and once again the 'katagenic' members appear to represent curtailment of the potential hypermorph which is itself steadily enlarging. These phenomena can also be stated in terms of paedomorphosis and

---

* It is now probable that Graptolites are related to the hemichordate Pterobranchiata; the more familiar terminology of the standard palaeontological texts is used here.

retardation (de Beer 1940*a*); the 'katagenic' phases may be regarded as neotenic modifications of enlarging thecal form.

One of the most controversial issues in palaeontology has centred upon the interpretation of ontogeny of ammonoids. It is a well-established fact that many

argenteus   priodon   cf. colonus

A   B   C
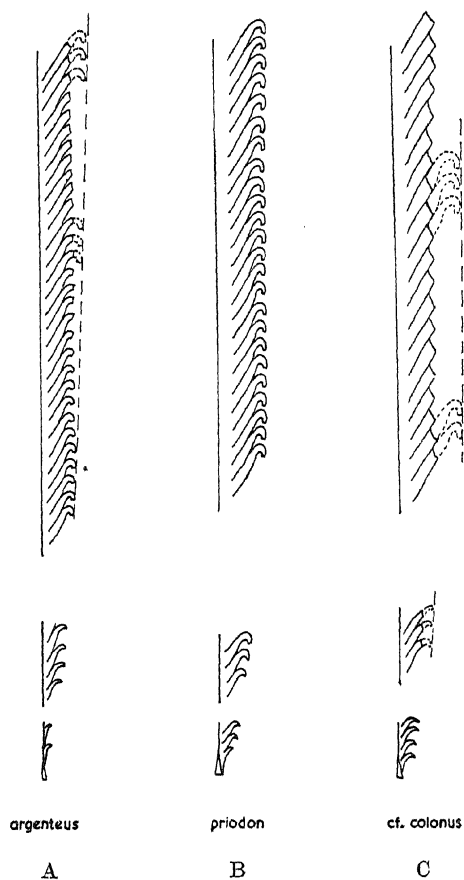
FIGURE 25. Application of the potential hypermorph concept to 'hooked' *Monograptus* (semi-diagrammatic). The most proximal, a few 'immature', and a range of fully-developed thecae are shown in each case; the suggested potential hypermorph developments are in dotted lines. A based on *M. argenteus*, B on *M. priodon*, C on *M. colonus*.

lineages of ammonoids show an apparent recapitulation, which has been variously interpreted. This is particularly true of lineages showing increasing size, and of certain Palaeozoic groups. It is often so thoroughgoing that the centre portions ('nucleus') of the Upper Permian ammonoid *Cyclolobus kullingi* from Greenland, for example, was placed in a new genus *Godhaabites*, and dated nearly one geological period too early, with curious stratigraphical results. But this 'recapitulation' is by no means universal, and Spath's account (1938) of the Liassic Liparoceratidae has given clear evidence of what has been called prospective development (Swinnerton 1938), the characters of *descendent* forms appearing in the earlier whorls. According to Spath the Liparoceratids were derived from *Tetraspidoceras*-like

forms by great inflation and increased involution of the outer whorls; the latest members of the lineage show progressive change until a nucleus with capricorn ornament enlarges to fill the entire ontogeny (paedomorphosis). It seems reasonable again to suppose that fairly late stages showing modest expansion of the outer whorls would, if they had been able to attain sufficient size, have given a *Liparoceras*-like hypermorph, larger than ancestral forms. Here again we may interpret a katagenic series as a neotenic or paedogenetic curtailment of increasingly large inflated forms. Similarly with pseudoceratites, 'katagenic' ammonoids with simplified suture lines; if these had grown to really large sizes the sutures would undoubtedly have become more complex and ammonitic. (It is interesting that some Liparoceratidae show the pseudoceratitic trend of suture lines.) Newell (1949) gives a most convincing figure showing the allomorphosis of suture length against conch diameter in Permian ammonoids, and no doubt heterauxesis curves would be similar.

Evolutionary changes in relative growth patterns

Some examples, taken from fossil vertebrates, may serve to indicate a further problem in which relative growth studies are essential.

Bystrow (1935, 1938, etc.) has made a detailed study of the skulls of various labyrinthodont Amphibia, and finds that zones of relatively intense growth, shown, for example, by extreme separation of snout tip and external nares, or nares and orbit, are often shown by 'streaking-out' of the ornamentation. Different genera in the same family show general similarity, but quite remarkable differences exist between different forms (cf. also Romer (1947), for an admirable recent general account of the group). Bystrow showed that the central parts of many of the dermal bones have a more or less non-directional ornament, and such areas can be marked out on a drawing or photograph, cut out, and reassembled to give a short, round-snouted, large-eyed 'skull'. Such a pattern is like that of the small Amphibia classed as Phyllospondyli. Romer (1939) figures many growth stages of various labyrinthodonts, and finds that all can be traced to more or less similar small forms. He has concluded that the Phyllospondyli are no more than very juvenile Labyrinthodontia, though Watson (1940) considers that the group has a real taxonomic value. It is, of course, quite likely that some Phyllospondyli were sexually mature neotenic forms; a larger labyrinthodont of similar skull shape, *Dwinosaurus* (Bystrow 1938) has been fully discussed as a neotenic form, and is probably a 'giant', with curtailed development history.

The excellent account and drawings of one of the early Triassic labyrinthodonts, *Benthosuchus sushkini*, by Bystrow & Efremov (1940), allow a detailed study of relative growth to be made. Preliminary study showed that the breadth of the skull table (neglecting the tabular horns) is a convenient dimension for comparison (x). The individuals used range from $x = 12\cdot8$ mm. to $x = 83$ mm., and very much larger individuals (up to $x = 200$ mm. +) are known in fragments. Over this range numerous measurements show very close fit to the simple allometry formula, $y = bx^{\alpha}$; direct arithmetical plottings show obvious deviations from straight lines in many characters, so thirty-two variables have been calculated in terms of the simple allometry formula. The smallest individual figured by Bystrow & Efremov

is taken as thoroughly typical of its size group. The fit of points over a range of $6\frac{1}{2}$ times in breadth and $8\frac{1}{3}$ times in width is very good, and larger less complete specimens (up to $x = c.\ 140$ cm.) fit the rectilinear log/log regressions very well. From the graphs, the proportions to be expected at $x = 5$ mm. and $x = 200$ mm. (the dimensions of the table in Bystrow & Efremov's suggested reconstructions of a 'very large' individual, their figure 62) have been determined, and give entirely self-consistent reconstructions (see figure 26). The large individual, as
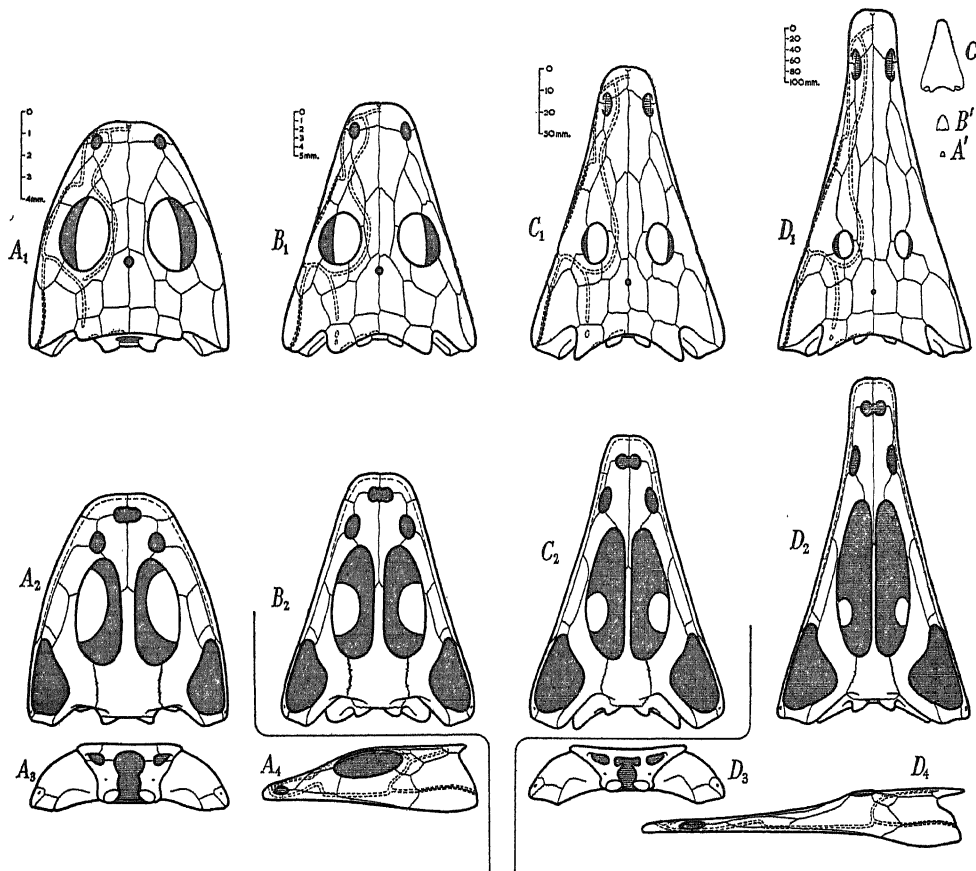
FIGURE 26. Growth in the skull of the labyrinthodont *Benthosuchus sushkini*; based on data from Bystrow & Efremov (1940) analyzed in terms of simple allometry. *A*, conjectural stage, skull-table (x) 5 mm. broad; *B*, $x = 12\cdot5$ mm. (cf. Bystrow & Efremov, figures 56–59); *C*, $x = 50$ mm. (cf. Bystrow & Efremov, figures 63–64); *D*, $x = 200$ mm. (see text). Suffixes: 1 = dorsal, 2 = palatal, 3 = occipital, 4 = lateral view. The true comparative sizes of the individuals are shown to the right, on the same scale as $D_1(A', B', C')$.

tentatively reconstructed by Bystrow & Efremov, is, according to their other figures of fragmentary large individuals, less elongated, and much broader at the snout, than it should be—the new reconstruction (figure $26D$) is much more in agreement with the material. The small individual (figure $26A$) is drawn without allowing for any structural modifications due to immaturity—e.g. it is quite possible that at such a size the parasphenoid and pterygoids were not suturally united. Nevertheless, there is much resemblance to phyllospondyls and to small

early members or collaterals of the approximate phylogenetic lineage of *Bentho-suchus*, such as *Eugyrinus*. Only two groups of measurements show marked deviation from simple allometric change. Medium-sized to large specimens ($x = 50$ to $x = 140$ mm.) show negative heterogony of tabular horn ($b = 0.55$, $\alpha = 0.80$) and depth of otic notch ($b = 0.34$, $\alpha = 0.98$); on this basis each should be c. 4·2 mm. long in the smallest skull figured by Bystrow & Efremov; the actual measurements are 2·4 and c. 2·6 respectively, indicating a different early development of these structures. It is probably significant that most phyllospondyls have a characteristic broad table and small tabular horns, with often a poorly defined otic notch. It would seem that an unusually simple growth pattern may exist in the skull of *Benthosuchus*. Preliminary investigation of other genera suggests that the same is true in other labyrinthodonts, and that many of them can be obtained from juvenile skulls, generally similar to that here deduced for a small *Benthosuchus*, by varying the parameters in the allometric equations for each linear measurement. In the case of the suggested development series *Branchiosaurus ambly-stomus*—*Pelosaurus*—*Onchiodon* (Romer 1939) there is a sharp break in the log/log plots of nearly all skull dimensions (y) plotted against breadth of table (x); *Branchiosaurus* has a very different conformation from the others; nevertheless, the trends of the regression lines for the *Branchiosaurus* members of this assemblage are directed towards the *Pelosaurus-Onchiodon* lines (data from Romer, Bulman & Whittard, Watson and specimens). Watson (1940) considers that *Branchiosaurus amblystomus* is not a juvenile *Onchiodon*, and this is quite possible; on the other hand, the changes from small to large individuals of *Branchiosaurus amblystomus* do trend towards the *Pelosaurus* condition. Further study is required to show whether the special characters of the skull in small individuals are explicable as larval modifications (external gills are well known in such specimens).

During these growth changes in *Benthosuchus* the dermal-bone pattern behaves as though it were plastic; however, the lacrimal bone, which in very young individuals probably extended from the orbit almost to the external nares (as in many primitive Labyrinthodonts), becomes more isolated from both.

Still more remarkable changes in proportion are found in the approximate lineages in the evolution of major groups. Thus in the series from Osteolepid Crossopterygii through *Elpistostege* and *Ichthyostega* to early labyrinthodonts and primitive reptiles a tremendous change in proportions has been shown by Westoll (1938, 1943) and Romer (1941). The post-parietal, a large element almost half the length of the skull in Osteolepid, becomes a narrow element largely on the occipital surface; post-parietal and parietal together decline from 75 % of skull length in Osteolepids to 11 % in Pelycosaurs; the orbit moves anteriorly, and the snout is very greatly lengthened. Similar changes affect brain case, cranium, palate and jaws. This differentiation in form will ultimately require, for its understanding, study of the growth of various stages. It is clear that apparently quite small changes in the proportions of a juvenile skull, even supposing the α-parameters of the allometric equations for each part to be unchanged, would produce considerably greater changes in fully grown individuals; much more forceful is the effect of, say, a 10 % change in the α-parameters.

A proportional change of similar type, investigated by Parrington & Westoll (1940), is found in the palate and basicranial region of advanced mammal-like reptiles (cynodonts) and mammals. The changes are essentially due to approximation of the brain case (behind the level of the basipterygoid processes) to the back of the snout region. The region corresponding to the 'pterygoid girder' of cynodonts is greatly shortened in mammals, with numerous interesting anatomical consequences. Mammalian embryos show very little sign of ontogenetic shortening in this region, but in early members of the group there were probably interesting changes in growth gradients during development.

There can be little doubt that heterauxesis, in successive phases of evolutionary differentiation of this kind, differed more or less strongly from lineage allomorphosis (if, indeed, increase in size was general throughout the transition) or from the phylogenetic sequence of changes in pattern. It still remains a most significant problem to determine the mode of incidence, during successive ontogenies, of the changes which ultimately are of major taxonomic importance.

The way in which an 'old' development sequence leading to a particular structural pattern is modified to produce a markedly different pattern may never be determinable in fossil forms, so that we may be forced to continue for a long time the analysis of 'adult' patterns. Particularly in the case of what Simpson (1944) has called quantum evolution, such analysis is nearly always fragmentary and partial; any growth studies on material representing these critical phases are sure to be valuable.

THE BIOLOGICAL FOUNDATIONS OF MEASUREMENTS OF GROWTH AND FORM

BY C. H. WADDINGTON, F.R.S., *Institute of Animal Genetics, Edinburgh*

Valid biological conclusions can only be drawn from mathematical comparisons of growing forms in so far as the objects compared are biologically similar. Physiological analysis should, therefore, be prior to mathematical treatment.

Measurement is never an end in itself. It is carried out as a step in the analysis of some causal situation in which we are interested. Biological objects are usually so complex that they possess an embarrassing number of characteristics capable of being measured, and if the results of mensuration are to be of any interest or value, it is essential to keep firmly in mind the objective in relation to which they are required. Similarly, mathematics provides an enormous range of possible methods which can be applied to bodies of numerical data, but again no good purpose is served by merely passing the figures through these various mills unless the techniques are appropriate to reveal relations of a kind which are relevant to some predetermined point of view. Mathematics, like mensuration, can be adequately used as a tool only if one knows what one wants to do with it.

There are, of course, as Medawar has pointed out this afternoon, many possible purposes which may lead one to employ mathematical methods of analyzing measurements. In a few cases the objective may be merely to elaborate a short-

hand method of summarizing the values of a certain measurement taken through-out a period of time, as in his example of a normative curve relating weight to age in man. For such purely descriptive purposes one can be content with an empirical curve, to which can be fitted, with any necessary degree of accuracy, an algebraical function of any form one finds most convenient. The uses of a formula derived in this way are very restricted; it can be used for extrapolation beyond the limits of the data only with the greatest caution, its algebraic form cannot be taken to reveal any essential causal structure in the phenomena in question, and it can be used to compare two different biological systems only in a strictly mathematical sense—comparisons which involve biological relationships, such as phylogenetic affinity, belong to a realm in which an empirical descriptive formula has no place.

It is, however, usually in connexion with problems of this kind that the biologist wishes to call in mathematical aid. Sometimes he is interested in the formal structure of the causal system underlying the growth of an isolated system. This is the case, for instance, in the study of the growth of an isolated biological system, such as a colony of bacteria or even of a whole organism. Some charac-teristic is then defined as the 'size', be it the wet or dry weight, the volume or something else; this characteristic is measured at successive intervals of time; and the problem is to obtain insight into the structure of the system on which its increase depends. One might hope that, if the measurements are fitted by an algebraic function, they will be found to exhibit features which make one type of function appear much more attractive for this purpose than any other; and that type of function might then be taken to express the formal properties of the causal system involved. Unfortunately, as we all know, this optimism has turned out to be unjustified. Growth measurements can usually be equally well fitted by any of a large number of different formulae. We are, therefore, driven to attempt to discover the nature of the causal system in some other way, and, reversing the above procedure, determine the form of the algebraic growth function on other grounds before fitting it to the measurements. This is, of course, the origin of the highly complicated formulae which have been sometimes suggested, and which Dr Stoll has mentioned. Although the fact that they may contain so many para-meters that they can be fitted to almost any set of data may make them trivial from the point of view of the mathematician, that does not by any means deprive them of biological interest; one has only to contemplate, for instance, the effect of pituitary hormones on the growth in weight of man to realize that the causal system in this instance is in fact a highly complex one, which would have to be expressed by a function containing very many variables.

The realization of the complexity of the growth of all but the most elementary isolated biological systems has largely removed the interest from attempts to discover the formal properties of growth by mathematical analysis. Moreover, there is a further difficulty over and above that due to the complexity of the system. The measurements of the size of a growing biological object are nearly always subject to considerable random errors; either because one and the same individual during its growth fluctuates from the norm, its growth being at some periods slightly inhibited, at others slightly accelerated; or because the measure-

ments at different intervals have to be taken on different individuals. The data, therefore, provide only a broad track down which the curve must be drawn, which again increases the likelihood that any of a number of distinct functions will be found to be equally satisfactory.

Mathematical treatment of growth data may, however, be attempted from rather a different point of view, namely, in order to facilitate the comparison of different biological objects, which may be different species or the successive stages in the development of a single species. In such comparisons, interest centres more on form, or spatial pattern, than on mere size, however that may be defined. The spatial pattern of a multicellular organism, or even of the important parts of an organism, is nearly always highly complex, so that it cannot usually be simply described by a mathematical function. Its complexity can, however, be reduced in various ways, all of which, it must be remembered, amount simply to neglecting those aspects of the form which it is too much trouble to take account of. One of the most drastic types of simplification is to consider the form as made up simply of two masses and to study the relation between their sizes. This is the procedure which leads to the description of animal form in terms of allometry or 'relative growth'. Rather less drastic simplification along similar lines gives us a description in terms of growth gradients. It is not intended here to discuss the technicalities of fitting the conventional allometric function to the data, and thus determining the allometric constants. Instead I wish to inquire for what purposes these formulations are used, and if they are adequate to serve them. These purposes usually, if not always, involve the derivation of biological relationships. Thus from the inspection of growth gradients in Crustacea, one may attempt to deduce a physiological principle, for instance, that a fast-growing appendage tends to cause an increase in the growth rate of nearby appendages; or one may regard the differences between the allometric constants in male and female organs as revealing something of the nature of the physiological difference between the sexes; or again, comparison of the constants may be invoked to establish phylogenetic relations. There is no doubt that by the use of allometric constants one can quickly and easily describe the phenomena involved in such comparisons. But they are described in quantitative terms, and if this ostensible quantification is to be taken seriously, all such arguments really presuppose that the allometric constants express real physiological entities—entities capable of being controlled by the gene changes involved in phylogenetic advance or the hormone differences involved in sex determination. Now this cannot, strictly speaking, be true. The formal properties of the allometry equation are such that if two different parts of the organism have an allometric relation to the whole, they cannot also have it to one another; the equation, therefore, cannot express a physiological relationship which holds generally between growing parts. Moreover, we know that in many cases the equation can only be fitted by invoking a series of sudden changes in the constants, so as to fit a number of straight lines to data which would more naturally be described by a curve.

This means that the allometry equation has the status, not of a physiological principle, but of a rough and ready shorthand method of description. The actual

biological changes which must have occurred when one species or genus of anteater evolved into another were not in fact alterations in the allometric growth partition between the head and snout, because this growth partition is not a real physiological variable, but merely a derivative function from an empirical growth description which is a reasonably good approximation. For the same reason, the precision which at first sight appears to be imported into the discussion by determining the numerical values of the constants is largely spurious. Biologists, therefore, have no need to allow themselves to be overwhelmed by the impressiveness of the mathematical treatment. But this is not to deny that, when the situation allows of fairly simple treatment in allometric terms, that formulation may provide descriptive vocabulary which allows of rather more precision than could easily be obtained in any other way.

For many purposes, however, it is inappropriate to simplify the form so drastically as to reduce it to a relation between two elements of size. A suggestion of a method of treating such cases was made by D'Arcy Thompson, when he put forward the idea of considering the form as inscribed in a deformable network of Cartesian co-ordinates. The method can, of course, only be used to compare two forms if these are homeomorphic, that is, related to one another in such a way that, if one shape were drawn on a deformable surface, the drawing could by suitable stretchings and contractions be made to resemble the other shape. One shape may be more 'complex' than another homeomorphic with it—for instance, if straight lines in one form become wavy lines in the other. But there is another way of defining complexity, which takes account only of the addition of new elements to a pattern, so that the degree of complexity must be measured in whole integral units. In the pure realm of geometry, such measurements of complexity belong to the realm of topology. The point that is being made throughout this discussion—that if mathematics are to be used in biology, it is essential to ensure that they are relevant and adequate in relation to some definite biological problem—is very well illustrated by the obvious irrelevance of classical topology to animal form. From the strictly geometrical point of view, each new isolated cavity appearing within a body, such as a blood island in an embryo, or a food vacuole in an amoeba, raises the degree of complexity of the form by one unit, whereas a new external excrescence, even if it is as functionally important as a limb, leaves the degree of complexity unchanged. It is obvious in this case that we cannot pass simply from the mathematical relations to statements about biological relations. We should require for that a specifically biological analogue of topology—a new branch of science in which relations of the logical kind used in normal topology were set up between entities whose definition was essentially biological. I have suggested (Waddington 1940b) that these biological definitions might be founded on a consideration of the kinds of spatial deformation occurring during embryogenesis, but this idea still awaits further elaboration.

So long as the shapes we wish to compare are homeomorphic, there is no doubt that it is theoretically possible to find a mathematical transformation which will convert the one shape into the other, and which may thus be taken to define their relationship. This transformation may be illustrated visually, as D'Arcy Thompson

did, by drawing systems of co-ordinate lines which connect homologous points on the two forms. It should also be possible to describe the transformation algebraically. Indeed, Medawar (1944, 1945) has made a beginning towards doing this for one case, that of the changing pattern of the human body during its development from the early foetus to the adult. He has, of course, drastically simplified the form. Whereas D'Arcy Thompson, in his diagrams, reduced three-dimensional forms to two-dimensional outlines by projection on to a plane, Medawar has gone one step farther and again projected certain salient points in his two-dimensional outlines on to a line, thus reducing his original three-dimensional form effectively to a one-dimensional sequence of points. As development proceeds, the intervals between these points alter; and it is, of course, possible to find empirical algebraic expressions to describe these alterations. Medawar has thus shown the way to produce a vocabulary which could be used to do for our discussions of relatively complex shapes what the allometry language does in relation to simple two-element forms. That is to say, it would enable us to give relatively precise descriptions of certain observations. But it would also suffer from the same limitation, that of being an empirical description with no simple relation to the effective causal system whose nature and modifications constitute the biological problem which we have to understand.

In order to illustrate what is meant by the difference between mathematical and biological relations, it may be well to consider the development of the wing of *Drosophila*. This organ is extremely simple, as biological organs go; a flat, two-dimensional plate consisting of two layers of cells closely apposed to one another, with a smooth, roughly oval, outline, and marked only by five longitudinal veins and two cross-veins. We know many modifications of the normal or wild type shape, some produced by mutant genes, others by a variety of experimental procedures. The relations between these various shapes would seem particularly appropriate for analysis by co-ordinate transformations. The wing is also very suitable for analysis by biological methods, since it is easy to study the development of the normal and mutant forms (Waddington 1940a), and Henke and his co-workers (see Henke 1947) have found that many specific types of modification can be produced by short treatments of the pupa with just sublethal temperatures. Moreover, each cell on the wing surface gives rise to a single hair, and by counting these it is simple, on the one hand, to discover something of the alterations in cell number and arrangement involved in the various modifications, and, on the other, to express the mathematical relations as transformations of the whole wing-space, rather than merely of the outline drawing.

Mere inspection of the developing wing shows that it takes place in a series of phases, in each of which some process, contributory to the change of size and form, is occurring; in each of these phases some characteristic modifications can be produced. The preliminary phase (or more properly, series of phases) occurs before the wing fold is everted from the imaginal bud. Some of the mutant genes which produce their effects at this time cause the pattern of veins to be shifted in relation to the epithelia; thus in *shifted* the longitudinal veins are pinched together, in mutants such as *four-jointed* they are spread apart. In both cases they are moved within

the general spatial system defined by the bulk of the epithelia; and if they were used, as it would be natural at first sight to use them, as fixed points around which to organize a system of co-ordinate lines, we should succeed in instituting comparisons which were quite false as far as the epithelia are concerned.

Shortly after the eversion of the wing blade, a first phase of increase in size occurs (figure 27). This is due mainly to an increase in cell number. The cell divisions appear to be oriented preponderantly in the longitudinal direction, and if they are partly inhibited by a sublethal temperature shock, the resulting wing is smaller and at the same time relatively broader than normal. Genetic factors may influence the orientation of the mitotic spindles, making the wing either longer (as in *lanceolate*) or broader (as in *broad*) without changing the overall size, or change both the predominant orientation and the number of divisions (as in *expanded*, which is both broader and larger than normal).

This phase of division is followed by one in which the size of the wing increases by the enlargement of its cells, accompanied by the accumulation of body fluid within the hollow wing-sac. After reaching a maximum of inflation, the wing begins to be reduced in size, owing to a contraction of the epithelia and the withdrawal of the fluid. Many genes are known which influence this process of contraction, either by exaggerating or inhibiting it. Temperature shocks at this period usually produce phenocopies of the well-known mutant type *dumpy*, in which the contraction is exaggerated.

Following this stage, the two-phase cycle is repeated. There is first a second period of cell division. In this case, the spindles are oriented mainly transverse to the length of the wing, and inhibitions due to temperature shock give elongated wings. There is a later phase of this period (see Waddington 1940a, text-figure 4) during which growth occurs mainly near the tip of the wing, and temperature shocks during this produce wings in which the tip is rounded. This phase is succeeded by the second period of cell expansion. According to Henke, this is at first mainly in length, then equally all round and finally in breadth. Inhibition by temperature shock at the beginning of the phase gives a second broad type, at the middle gives a smaller form isomorphous with the wild, and towards the end an elongated form. The genes known to operate at this time give either an isomorphous decrease (e.g. *miniature*) or increase (*venae abnormae*). Finally, the last stages of the expansion occur earlier on the under-side of the wing than on the upper, and temperature shocks may cause the blade of the wing to be warped upwards or downwards out of the flat plane.

What has been given above is, as might be expected, a considerably simplified account of the causal system which determines the shape of the wing. There are actually certain interactions between different phases; for instance, elongated wings respond to the forces of contraction in a slightly different pattern to that of rounder ones. But enough has been said to show that there are many factors involved in wing form which are not taken account of in a straightforward mathematical description of the end result. Yet these factors are extremely important if we wish to draw any biological conclusions from the comparison of different wing forms. We may, for instance, have two rather similar broad-winged species,

1

2

3

4

5

6

7

8

9

a

c

g

i
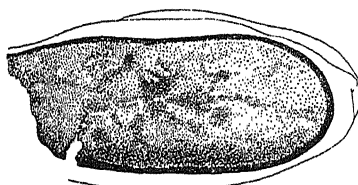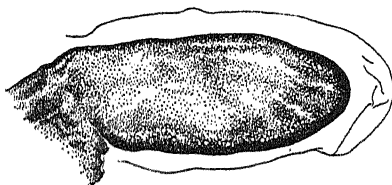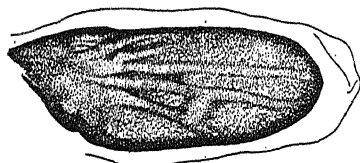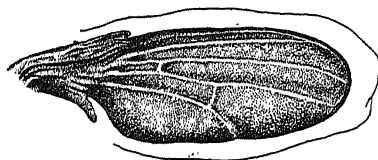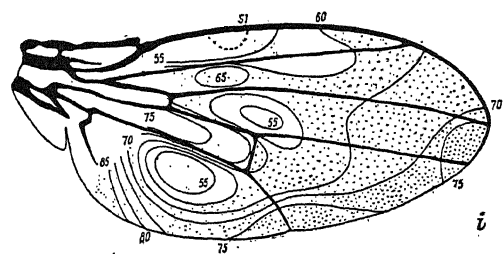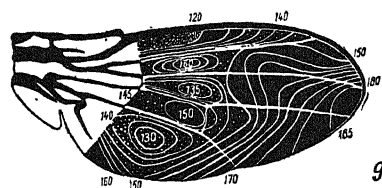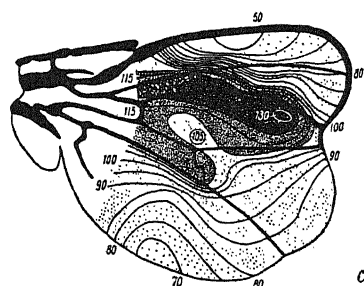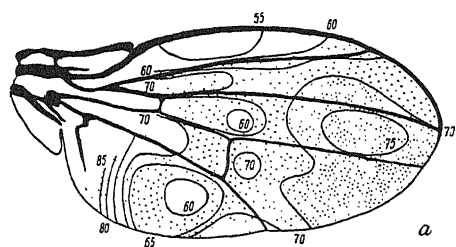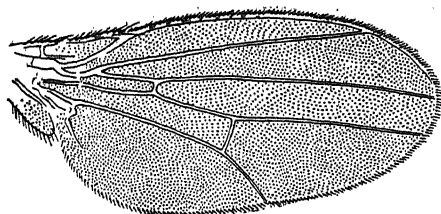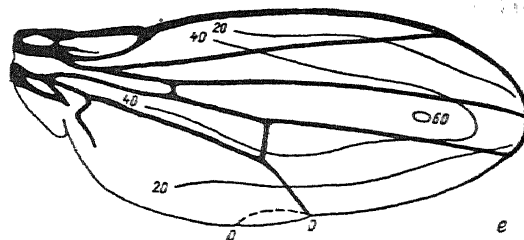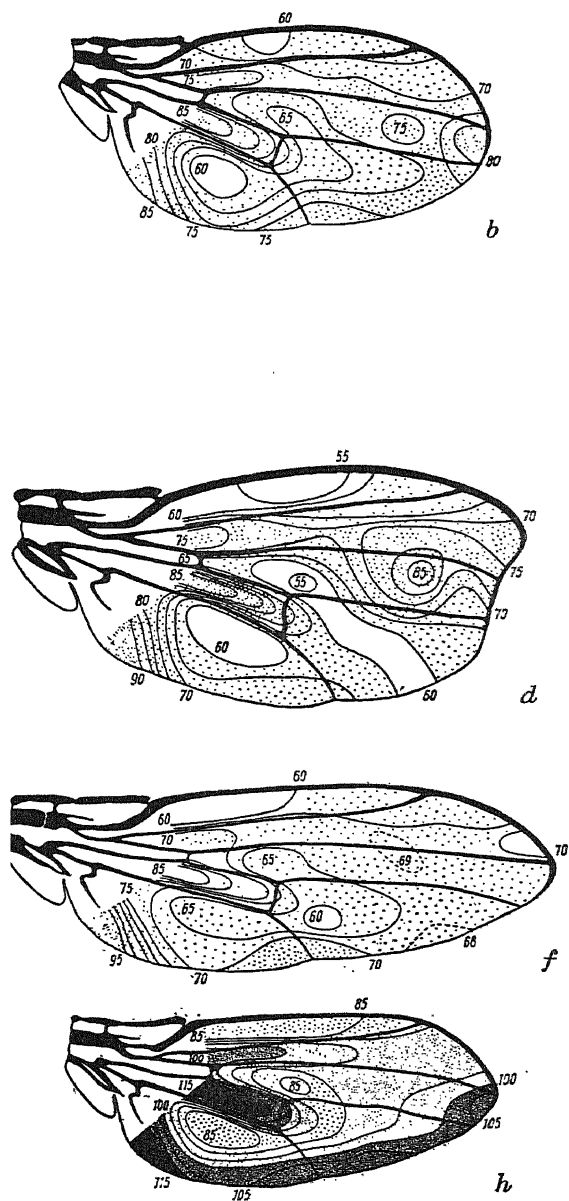
FIGURE 27. Growth and form of the wing in *Droso-phila*. On the left side are drawings of the wing at successive stages of its development from the time of puparium formation to the formation of the imago: stages 2 to 4, first period of cell division; 4 to 6, first expansion and contraction; 6 to 8, second period of cell division; 8 to 9, second expansion. (After Waddington.)

On the right are outline drawings of normal wing, *i*, and modifications produced by genetic factors (*a, c, g*) or by temperature shocks (*b, d, f, h*) on which have been drawn contour lines of equal cell-density. *a* broad, *b* phenocopy of broad, both produced by effects on first period of cell division. *c* dumpy, *d* phenocopy, produced by effects on first expansion and contraction; *e* shows distribution of differences in cell-density between *e* and a normal wing, *i*. *f* elongated modification produced by temperature shock at second period of cell-division. *g* miniature, and *h*, a phenocopy of it produced by effects on middle period of second expansion. (After Henke.)

both apparently derived from some other species with a more elongated wing; but we should have to assess the difference between the broad-winged forms quite differently according to whether they have both been produced by modifications of the same phase of development or alternatively owed their characteristic features to entirely unlike processes.

We can, in fact, use mathematical techniques for assessing comparisons only when, and in so far as, we are willing to consider the phenomena under discussion as biologically comparable. Only if we are dealing with modifications of one single developmental process is it justifiable to omit all other considerations and discuss a biological relationship in purely mathematical terms. Thus nothing of importance is omitted by a geometrical description of the various shapes produced by different degrees of contraction at the end of the second phase described above, or by changes in the predominant orientation of the division spindles in the first phase. As regards the former, Henke has in fact already made the first step in the analysis, by plotting a diagram which shows the local differences in cell densities between normal and *dumpy* wings. This is another way of expressing a deformation of a co-ordinate net, and could also be put into algebraic terms, if one wished, and thus reduced to forms capable of numerical estimation—but as yet no occasion for quantitative statements has arisen in this connexion.

There may be some temptation to dismiss the thesis which has been urged above as a counsel of perfection. Certainly there are many cases in which we wish to compare biological shapes about whose developmental physiology we have little or no information. But to acknowledge this is no more than to admit the essential and never-to-be-overcome incompleteness of science. It would undoubtedly be foolish to refuse to employ the techniques of mathematical analysis until one were certain that the processes under consideration were truly comparable from a biological point of view, because such certainty can never be finally achieved; even within, say, the first period of cell division in the *Drosophila* wing, the processes leading to increase in breadth are in some way different from those causing elongation. But even so, it remains true that the first essential is to push the biological analysis as far as it can be carried. The validity of any biological conclusions which may be drawn from measurements of size and form depends far more on the adequacy of the physiological insight on which they are based than on the precision of the mathematical techniques used to summarize and compare them.

GENETICAL ASPECTS OF SIZE ALLOMETRY

By E. C. R. REEVE

Professor Haldane has described a simple and improved method of estimating differential growth ratios, which I believe he and Teissier developed independently. This method avoids the bias arising in regression estimates when both dimensions being compared are subject to errors, and will be of great practical value, particularly if the standard errors of the growth coefficients can be easily compared.

Haldane has suggested that the use of this method might lead to different con-
clusions in my study of differential growth in the skulls of anteaters (Reeve 1940)
which was illustrated in Professor Young's paper to-day. The growth coefficients
(α) of maxilla against cranium length, calculated by the regression on logarithms
and by Haldane's method are given in table 3 for the three genera of anteaters.

TABLE 3

	values of α estimated by	
genus	regression formula	Haldane's formula
Myrmecophaga	$1\cdot77 \pm 0\cdot14$	$1\cdot88$
Tamandua	$1\cdot36 \pm 0\cdot08$	$1\cdot52$
Cyclopes	$1\cdot26 \pm 0\cdot11$	$1\cdot41$

Haldane's formula gives higher estimates than the regression method, but the
differences are small because of the high correlation between the two dimensions
in each genus. Both formulae evidently lead to the same conclusion, namely that
Myrmecophaga has a higher relative growth ratio of the snout than the two
smaller genera, which do not differ much in this respect. This is of interest
because, on the basis of general appearance and habits, one would naturally class
Tamandua with *Myrmecophaga* and not with *Cyclopes*.

The geneticist who is interested in the factors causing variations in proportions
takes a different point of view from those discussed earlier to-day. This may be
illustrated from a study of the inheritance of body size in *Drosophila melanogaster*,
in which Dr F. W. Robertson and I are collaborating at Edinburgh. In this animal
we are, of course, only concerned with adult size allometry (allomorphosis), since
there is complete metamorphosis. A number of lines has been selected during many
generations for large or small body size, using in some cases wing length, in others
thorax length as the measure of size. In the unselected stock reared under the
usual conditions at 25° C, variations in size and proportions are caused by both
genetic and environmental factors, whose net result is that both dimensions tend
to change in the same proportion—there is a relative growth coefficient of about 1.
The ratio between the two dimensions is, however, subject to environmental
modification, since both size and proportions are altered by change of temperature,
and may be altered by qualitative changes in diet, though reduction in quantity
of food tends to reduce size without altering proportions.

Genetic changes in size and proportions can also be brought about by continued
selection of one or the other dimension, as indicated by the results in table 4 of
selecting lines taken from a single wild stock for long and short wings and thorax
respectively.

The ratio of the percentage changes in the two dimensions (unselected/selected),
given in the fourth column, may be described as a coefficient of genetic allometry,
and this evidently varies with the direction and method of selection. Different
values may also be expected in selecting from different stocks.

A further complication occurred in the line selected for long wings. This line did
not become genetically homogeneous, even after seventy generations of selection,

although size increased little during much of this period. This is clear from the fact that as soon as selection for long wings is relaxed size declines until wing length has lost about half its total advance. But thorax length only loses about 20 % of its advance, so that the ratio of the percentage changes under relaxed selection is about 2·5. In other words, the coefficient of genetic allometry is 1·0 during increase in size brought about by selection for long wings, but is 2·5 during the reduction in size which occurs when selection is relaxed.

<div align="center">

TABLE 4

Changes in size and proportions caused by selection at 25° C

</div>

Nettlebed stock line selected for	percentage changes in		ratio of percentage changes	genetic correlation
	wing length	thorax length		
long wings	+ 7·2	+ 7·3	1·0⎫	
long thorax	+ 5·4	+10·8	0·5⎭	0·73
short wings	−23·5	−11·4	0·49⎫	
short thorax	− 9·0	− 8·2	1·1 ⎭	0·71
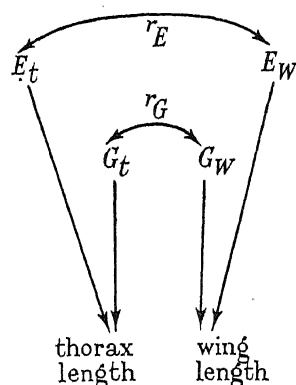
FIGURE 28

These results can only be understood in terms of the causal relationships underlying the variations in proportions. The wing length of any individual can be thought of as the resultant of the sum of the effects of genetic factors affecting wing length (the 'genotype' for wing length, G_W) and of the environmental factors affecting wing length (the 'environment' for wing length, E_W). In the same sense we can define G_T and E_T, the genotype and 'environment' for thorax length. Neither set of factors will have identical effects on the two characters in different individuals, but there will generally be a correlation between the genetic factors (r_G) and a correlation between the environmental factors (r_E), so that either genetic or environmental changes in one character will tend to produce corresponding changes in the other. This method of approach was developed by Wright (1921), and has been applied to livestock in the calculation of appropriate selection indices based on several characters (Hazel 1943). The causal relations affecting wing and thorax length are best set out in a path coefficient diagram (figure 28).

In this diagram the two-headed arrows indicate correlations between the causal agencies E_T, E_W, etc., and the single-headed arrows indicate paths of determination

from these to wing and thorax length. Wright developed this method of diagrammatic analysis and also methods of attaching coefficients to the different paths so that the relative importance of each factor could be estimated. Estimates of this kind can only be made on the basis of certain assumptions regarding the nature of the genetic and environmental factors involved. The above diagram, for instance, assumes that there is no correlation between G_W and E_W, or between G_T and E_T.

Clearly, in a genetically uniform stock, variations in size and proportions would depend entirely on the environmental chain of causes ($E_T.r_E.E_W$), while both causal chains will affect the allometry coefficient in an unselected stock. Changes in mean size and proportions due to selection will depend entirely on the genetic causal chain, and it is evident that the change in thorax length when wing length is selected (or vice versa) will depend on the size of r_G and the relative magnitude of the genetic variations in the two characters. In fact, the coefficients of genetic allometry, given in the last column of table 4, are really genetic regression coefficients of the unselected on the selected dimension, and the two genetic regressions obtained when we select for long wing and long thorax respectively will not be the same unless the genetic correlation $r_G = 1$, i.e. when all genes affecting either dimension have a proportionate effect on the other.

By analogy with linear regression methods, we can obtain an estimate of the genetic correlation as the geometric mean of the two coefficients of genetic allometry (i.e. of genetic regression) found when we select for long wing and long thorax (or short wing and short thorax). These estimates are given in the last column of table 4. The two values are 0·73 on selection for large size and 0·71 on selection for small size, which show unexpectedly good agreement. This implies that about 70 % (r_G) of the genetic variance of each character is common to both of them. Methods of estimation based on progeny tests give rather higher values of r_G, and the problem requires further study. The behaviour of the long-wing line when selection pressure is relaxed introduces further complications, which will be examined in detail in a forthcoming paper.

The method of analysis discussed here must be applied with caution, since it involves hypothesis about the nature of genetic and environmental variations which need testing, and further experimental work is particularly needed. It is likely that both genetic and environmental changes will be found to alter the genetic and environmental correlations, so that predictions based on one stock or set of conditions may not apply to others. But this approach does enable us to broaden the scope of allometric studies, and brings to light the complexity underlying the apparent simplicity of an allometric-size relationship.

THE EXACT DESCRIPTION OF LIVING ORGANIZATION

By J. Z. YOUNG, F.R.S., *Department of Anatomy, University College, London*

The papers contributed to the discussion give an idea of the large amount of work that has been devoted to the study of the relative growth of parts and the determination of the shape of the body. There is a danger that the fascination of such

investigation may lead to the prosecution of studies of this sort without sufficient consideration of their objective. Why do we want to study the shapes of living organisms? Several replies no doubt spring to mind, for instance, that we make such studies in order to facilitate comparisons between organisms. This, however, presupposes that the sort of comparison that can most usefully be made is essentially a visual one, involving the geometric outlines of the parts, a view that can only be accepted with reservations and if its implications are understood.

For some centuries the science of anatomy has grown up by the study of the visible outlines of the parts of the body. In the case of the anatomy of man, the obvious value of such studies, especially for the surgeon, has encouraged their pursuit into the most minute details. Valuable as such work undoubtedly is, biologists have long realized that it has severe disadvantages. However important it may be to know the details of the shape of any organ or part, we are not, as biologists, ultimately interested in these shapes but in the life of the organism. In the last resort any biological study must deal with the organism as a self-preserving system—we are interested in whether and how each individual or population will continue to live. Study of the visible shape is an essential part of almost every biological study, but it is only a part. The anatomist has long ago discovered that if he is to make his science useful he must treat the shape of each structure as a reflexion of the processes going on within it. The anatomist, microscopist and biochemist can thus pursue together the study of the organism as a steady-state system, discovering the genetic, internal and external factors that maintain its balance.

It is important that no single approach to biological problems should obtain an undue ascendancy; in particular, that the study of shape should not outstrip that of process, as it certainly did in the last century, producing a somewhat sterile form of anatomy. There is a danger that the attractions of the mathematical study of form may lead again to an undue emphasis on this single facet of life. We must avoid reviving the view that study of form in some way gives us a specially clear idea of what the organism is 'really like'. For this reason I suggest that it is by no means improper to inquire what it is that we hope to gain from study of the shapes of organisms.

It may well be that we shall find that by forcing ourselves to attend to wider aspects of biological process we may discover new fields that are amenable to the sorts of exact treatment that are under discussion. Medawar has shown that it is possible to give a quantitative specification of the changes in shape of homeomorphic objects. Waddington, however, has indicated that severe difficulties arise in the practical applications of these principles, and in any case they do not enable us to give anything approaching a complete definition of the form of the parts of organisms, such as would allow us to make interesting comparisons between them.

In evolutionary studies we are interested to trace the changes in living organization that enable organisms to enter new environments. In many cases evolutionary 'progress' consists in the invasion of environments that involve an increasingly large difference between the composition of the organism and its surroundings, and hence necessitate more and more complex devices to ensure the maintenance

of the organization (Young 1938). Can methods of exact description be used to
make statements of this sort more precise? Is there any way in which we can
usefully assess the degree of complexity of the organization of an animal or plant?

As an example of the problem involved one may consider the case of the
epidermis. In *Amphioxus* this is composed of a single layer of cells, all of one type,
and this provides a suitable boundary layer for a marine organism, whose internal
osmotic concentration differs little from that of the surroundings. In later aquatic
forms, however, for instance, the lamprey, the epidermis contains many layers,
the outer ones keratinized, and in addition several other types of cell, including
gland cells. This many-layered skin plays a large part in maintaining the difference
between the fish and its environment; for instance, in fresh water, the secretion
of mucus assists in the osmoregulation. In mammals the epidermis becomes still
thicker and more keratinized and produces a considerable variety of types of cell,
concerned with maintenance of the shape of the animal, waterproofing it and
keeping it at constant temperature and free of infection, these being among the
most important factors that enable the mammal to exist as a watery system in
a dry medium.

There is clearly a sense in which the skin of mammals has a more complicated
organization than that of *Amphioxus*, and we can specify this quantitatively by
saying that there is only one type of cell in the earlier animal but at least five sets
in any one part in the latter (cells of the Malpighian layer and their keratinized
descendants, sweat glands, sebaceous glands, hair follicles, dendritic cells). But
this enumeration only gives a faint idea of the conditions in the two animal types.
The difference would appear still more marked if we consider the various 'species'
of epidermis to be found in different parts of the body, for example, corneal skin,
nail, skin of the soles, and so on. Billingham & Medawar (1949) have recently
emphasized that these are self-reproducing 'genetic' entities, determined during
development. In order to specify them properly we need something more than
a definition of the shapes of their cells, however exactly given. We need knowledge
of the underlying biochemical differences that have developed between them and
then, in turn, of the biochemical hereditary mechanism that has produced these
differences. This may seem so obvious as to be a truism. Indeed, it has been
a truism now for twenty years or more that understanding of such problems lies
in analysis of their biochemical background. Nevertheless, to-day, although we
spend much time on the mathematics of 'form', there are still few who feel the
urgency of the need to link the form with its biochemical determinants. Surely it
is more important that we should find ways of doing this than that we should
refine our methods for definition of the form. It is plausible to argue that exact
definition is necessary before the determinants can be sought, but the experience
of most biologists is that the first requirement in such an investigation is to find
material amenable to experiment, and that the mathematical tools will need to
be forged to fit the particular situation that is chosen.

The problem is to find methods for the comparative study of morphogenetic
systems. It is curious that in spite of the comparative traditions that are so deeply
embedded in the science of embryology there has been rather little detailed com-

parison of developmental processes in related forms, directed to discovering the nature of the processes by analysis of the differences between them. Something has been done in related amphibians and species of *Drosophila*, but we still have no knowledge about the way in which, for example, uniformly rhythmical metameric systems, such as those of the vertebrae or the teeth, become broken up into a series of diverse parts. Further close study of the development of the backbone, and especially the teeth, of various vertebrates might show how the complicated and varied forms are produced from simpler ones.

Study of the differences between the shape of one part alone is liable, however, to detract attention from other differences between organisms. For example, the careful work of Reeve (1940) showed that the skulls of the three genera of xenarthrous anteaters differ only slightly in the relative growth rates of the maxilla and the skull as a whole. The enormously long snout of the great anteater (*Myrmecophaga*) is produced by a relative growth rate little higher than that of the shorter-faced tree-anteaters (*Tamandua* and *Cyclopes*), but the maxilla continues at its relatively high growth rate for much longer in the larger animal. One might therefore conclude that there was little genetic difference between these three anteaters except in the factors that control size. Inspection of the whole animals shows, of course, that however similar the skull growth rates may be other features are very different. *Myrmecophaga* has a long, bushy tail, the others have prehensile ones. The fur is long, dark and coarse in *Myrmecophaga*, short, soft and yellow in the others. There are, in fact, most profound differences throughout the whole life systems of the three animals. If our object is to compare these life systems we must have some way of studying the degree of correlation between the relative rates of development of various characteristics. At present we can do this only roughly (though still quite effectively) by listing the various appearances and modes of activity and behaviour. It would be an enormous advance if we could proceed to refer these differences to underlying differences of a chemical nature in the life processes of the populations, especially as epitomized during hereditary transmission and development.

The ideal is both obvious and utopian, but it can do no harm to keep it well in the foreground throughout studies of growth and form. Such studies are only part of the business of the exact definition of the lives of organisms as active self-maintaining systems, which is the central task of general biology.

REFERENCES

Aitken, W. G. & McKerrow, W. S. 1948 *Geol. Mag.* **85**, 19.
Alkins, W. E. 1920 *Manchr Mem.* **64**, no. II.
Alkins, W. E. 1923 *Manchr Mem.* **67**, 109.
Appleton, A. B. 1929 *C.R. Ass. Anat. 24th Meeting Bordeaux*, p. 3.
Arkell, W. J. 1933 *Proc. Cotteswold Nat. Field Cl.* **25**, 1.
Army Medical Department Reports (for years 1860–1913).
Beddoe, J. 1870 *Mem. Anthrop. Soc. Lond.* **3**, 384.
de Beer, G. R. 1940a In *The new systematics*, p. 365, ed. J. S. Huxley. Oxford University Press.
de Beer, G. R. 1940b *Embryos and ancestors.* Oxford: Clarendon Press.
Billingham, R. E. & Medawar, P. B. 1949 *Heredity*, **2**, 29.

British Association: Committee on Biological Measurements 1935 *Biological measurements*. London.

Brummelkamp, R. 1939 *Acta Ned. Morph.* **2**, 188, 260–268.

Bulman, O. M. B. 1933 *Biol. Rev.* **8**, 311.

Bystrow, A. P. 1935 *Acta zool., Stockh.*, **16**, 65.

Bystrow, A. P. 1938 *Acta zool., Stockh.*, **19**, 209.

Bystrow, A. P. 1942 *Acta zool., Stockh.*, **23**, 263.

Bystrow, A. P. & Efremov, J. 1940 *Trav. Inst. Paléontol. Acad. Sci. U.R.S.S.* **10**, livr. 1.

Camp, C. L. & Smith, N. 1942 *Mem. Univ. Calif.* **13**, 69.

Day, H. 1915 *Manchr Mem.* **59**, no. 4.

Denison, R. H. 1947 *Amer. J. Sci.* **245**, 337.

Dubois, E. 1897 *Bull. Soc. Anthrop., Paris*, **8** (ser. IV), 337.

Elderton, E. M. & Moul, M. 1928 *Ann. Eugen., Lond.*, **3**, 277.

Gray, S. W. 1946 *Amer. J. Sci.* **244**, 792.

Haldane, J. B. S. 1932 *The causes of evolution.* London: Longmans, Green.

Haldane, J. B. S. 1950 *Evolution*, **3**, 51.

Hammond, J. 1932 *Growth and the development of mutton qualities in the sheep.* Edinburgh: Oliver & Boyd.

Hammond, J. 1940 *Farm animals.* London: Edward Arnold.

Hammond, J. 1943 *Proc. Nutrit. Soc.* **2**, 8.

Hammond, J. 1947 *Biol. Rev.* **22**, 195.

Hartley, H. O. 1948 *Biometrika*, **35**, 33.

Hazel, L. N. 1943 *Genetics*, **28**, 476.

Henke, K. 1947 *Naturwissenschaften*, **34**, 149.

Hersh, A. H. 1934 *Amer. Nat.* **68**, 537.

Huxley, J. S. 1932 *Problems of relative growth.* London: Methuen.

Huxley, J. S., Needham, J. & Lerner, I. M. 1941 *Nature*, **148**, 225.

Jackson, C. M. 1915 *Morris's Human Anatomy.* London.

Jarvik, E. 1948 *K. Svensk. Vetensk.Akad. Handl.* 3e ser. **25**, no. 1.

Jarvik, E. 1950 *Medd. om Grönland*, **96**, nr. 4.

Jepsen. G. L., Simpson, G. G. & Mayr, E. 1949 *Genetics, paleontology and evolution.* Princeton Univ. Press.

Kemsley, W. F. F. 1950 *Ann. Eugen., Lond.*, **15**, 161.

Kermack, K. A. & Haldane, J. B. S. 1950 *Biometrika*, **37**, 30,

Lapicque, L. 1898 *C.R. Soc. Biol., Paris*, 5⁻, 62.

Lull, R. S. & Gray, S. W. 1949 *Amer. J. Sci.* **247**, 492.

Lumer, H. 1940 *Amer. Nat.* **74**, 439.

Ma, Ting Ying H. 1934 *Sci. Rep. Tohoku Imp. Univ.* ser. 2 (Geol.), **16**, 165.

Ma, Ting Ying H. 1937a *Palaeont. Sinica*, ser. B, **2**, f. 3.

Ma, Ting Ying H. 1937b *Palaeont. Sinica*, ser. B, **16**, f. 1.

Ma, Ting Ying H. 1938 *Bull. Geol. Soc. China*, **18**, 349.

Ma, Ting Ying H. 1940 *Bull. Geol. Soc. China*, **20**, 343.

Ma, Ting Ying H. 1943 *Research on the past climate and continental drift.* Vols. 1 (Ordovician), 2 (Silurian), 3 (Devonian). Yungan, Fukien, China.

Maclennan, R. M. & Trueman, A. E. 1942 *Proc. Roy. Soc. Edinb.* B, **61**, 211.

Martin, W. J. 1949 *Mem. Med. Res. Coun.* **20**.

McMeekan, C. P. 1940–41 *J. Agric. Sci.* **30**, 276, 387, 511; **31**, 1.

Medawar, P. B. 1944 *Proc. Roy. Soc.* B, **132**, 133.

Medawar, P. B. 1945 In *Essays on growth and form presented to D'Arcy Wentworth Thompson.* Oxford: Clarendon Press.

Ministry of National Service (1920). *Report* (1917–19) *upon the physical examination of men of military age by National Service Medical Boards*, **1**.

Needham, A. E. 1950 *Proc. Roy. Soc.* B, **137**, 115.

Newell, N. D. 1949 *Evolution*, **3**, 103.

Osborn, H. F. 1929 The Titanotheres of ancient Wyoming, Dakota and Nebraska. 2 vols: *U.S. Geol. Surv. Monogr.* no. 55.

Parrington, F. R. & Westoll, T. S. 1940 *Phil. Trans.* B, **230**, 305.

Pearson, K. 1897 *Proc. Roy. Soc.* **60**, 489.

Phleger, F. B., Jr. 1940 *Amer. J. Sci.* **238**, 643.

Phleger, F. B., Jr. & Putnam, W. S. 1942 *Amer. J. Sci.* **240**, 547.

Raw, F. 1925 *Quart. J. Geol. Soc. Lond.* **81**, 223.

Reeve, E. C. R. 1940 *Proc. Zool. Soc. Lond.* A, **110**, 47.

Reeve, E. C. R. 1941 *Proc. Zool. Soc. Lond.* A, **111**, 279.

Reeve, E. C. R. & Murray, P. D. F. 1942 *Nature*, **150**, 402.

Reeve, E. C. R. & Huxley, J. S. 1945 In *Essays on growth and form presented to D'Arcy Wentworth Thompson.* Oxford: Clarendon Press.

Richards, O. W. & Kavanagh, A. J. 1945 In *Essays on growth and form presented to D'Arcy Wentworth Thompson.* Oxford: Clarendon Press.

Robb, R. C. 1935 *J. Genet.* **31**, 39, 47.

Romer, A. S. 1939 *Amer. J. Sci.* **237**, 748.

Romer, A. S. 1941 *J. Morph.* **69**, 141.

Romer, A. S. 1947 *Bull. Mus. Comp. Zool. Harv.* **99**, 1.

Romer, A. S. 1948 In *Robert Broom Commemorative Volume*, Roy. Soc. South Africa, Cape Town, p. 45.

Romer, A. S. 1949 In Jepsen *et al.* 1949, p. 103.

Romer, A. S. & Price, L. I. 1940 *Spec. Pap. Geol. Soc. Amer.* no. **28**.

Ruger, H. A. & Stoessiger, B. 1927 *Ann. Eugen., Lond.*, **2**, 76.

Shepherd, R. H., Sholl, D. A. & Vizoso, A. 1949 *J. Anat., Lond.*, **83**, 296.

Sholl, D.A. 1948 *Proc. Roy. Soc.* B, **135**, 243.

Simpson, B. 1933 *Naturalist*, No. 914, 57 and 76.

Simpson, G. G. 1944 *Tempo and mode in evolution.* New York: Columbia University Press.

Snell 1891 *Arch. Psychiat. Nervenkr.* **23**, 436.

Spath, L. S. 1933 *Biol. Rev.* **8**, 418.

Spath, L. S. 1938 *The Ammonites of the family Liparoceratidea.* London: Brit. Mus. (Nat. Hist.)

Stirton, R. A. 1940 *Univ. Calif. Publ., Bull. Dep. Geol. Sci.* **25**, 165.

Størmer, L. 1942 *Norsk Geol. Tidsskr.* **21**, 49.

Stubblefield, C. J. 1926 *J. Linn. Soc. (Zool.)*, **36**, 345.

Swinnerton, H. H. 1938 Development and evolution. *Rep. Brit. Ass., Cambridge*, **108**, 57.

Swinnerton, H. H. 1939 Palaeontology and the mechanics of evolution. *Quart. J. Geol. Soc. Lond.* **95**, xxxiii.

Swinnerton, H. H. 1940 On the study of variation in fossils. *Quart. J. Geol. Soc. Lond.* **96**, lxxvii.

Teissier, G. 1948 *Biometrics*, **4**, 14.

Thompson, D'Arcy 1916, 1942 *On growth and form.* Cambridge University Press.

Trechmann, C. T. 1925 *Proc. Geol. Assoc.* **36**, 135.

Trueman, A. E. 1922 *Geol. Mag.* **59**, 258.

Trueman, A. E. 1930 *Biol. Rev.* **5**, 307.

Trueman, A. E. 1939 *Trans. Geol. Soc. Glasg.* **20**, 82.

Waddington, C. H. 1940*a J. Genet.* **41**, 75.

Waddington, C. H. 1940*b Organisers and genes.* Cambridge University Press.

Wallace, L. R. 1948 *J. Agric. Sci.* **38**, 93, 243, 367.

Watson, D. M. S. 1940 *Trans. Roy. Soc. Edinb.* **60**, 195.

Watson, D. M. S. 1949 In Jepsen *et al.* 1949, p. 45.

Weiss, P. 1949 In *The chemistry and physiology of growth.* Princeton University Press.

Westoll, T. S. 1936 *Geol. Mag.* **73**, 157.

Westoll, T. S. 1938 *Nature*, **141**, 127.

Westoll, T. S. 1943 *Biol. Rev.* **18**, 78.

Westoll, T. S. 1945 *Trans. Roy. Soc. Edinb.* **61**, 341.

Whittington, H. B. 1940 *Amer. J. Sci.* **238**, 241.

Wright, S. 1921 *Genetics*, **6**, 111.

Young, J. Z. 1938 The evolution of the nervous system and of the relationship of organism and environment. In *Evolution. Essays presented to E. S. Goodrich*, edited G. R. de Beer. Oxford: Clarendon Press.

Young, J. Z. 1945 In *Essays on growth and form presented to D'Arcy Wentworth Thompson.* Oxford: Clarendon Press.

Yule, G. U. & Kendall, M. G. 1945 *An introduction to the theory of statistics*, 13th ed. London: Griffin.

Oxidation-reduction processes in cultures of bacteria
I. The reducing power of *Bact. lactis aerogenes*

By A. C. Baskett* and Sir Cyril Hinshelwood, F.R.S.

Physical Chemistry Laboratory, University of Oxford

Cultures of *Bact. lactis aerogenes* show a reducing power towards oxygen (as measured by a modification of the Thunberg methylene-blue test) which is a rather stable characteristic of the growth under specified conditions.

During adaptation of cells to utilize a given carbon substrate the reducing power and the growth rate increase in parallel (over considerable ranges) to their respective optima.

In general the reducing power of fully adapted cultures is nearly constant for a considerable variety of carbon sources (12 to 15 units) in spite of a nearly fourfold variation of optimal growth rates.

The regularity is masked with certain substrates by abnormal enhancements or inhibitions occurring at high concentrations but reappears when the measurements are made in the low concentration range.

The result suggests that there is normally a rate-limiting step in the part of the cell mechanism which consumes molecular oxygen. This is not altered by adaptation to a new carbon source, though when the cells are first transferred to a new medium the proportions of various enzymes are unsuitable for the optimum utilization of the substrate, and the amount of oxygen which can be consumed is far below the maximum. Until these proportions are finally readjusted growth rate and the degree to which the oxidizing mechanism can be used increase in parallel, a limit to the adaptive process being set by the relatively unchanging maximum oxygen uptake.

Introduction

It has been previously established that during the logarithmic growth-phase cultures of *Bact. lactis aerogenes* possess a constant, characteristic reducing power (Lewis & Hinshelwood 1948). A number of authors, notably Thunberg (1920) and Quastel & Whetham (1924, 1925), have described a variety of techniques for measuring dehydrogenase activity by means of the reduction of methylene blue. A measure of enzymatic reducing activity is conveniently defined as

$$R = 10^4/tn,$$

where t is the time in minutes taken by a culture to reduce 0·73 m.-equiv. of dissolved oxygen together with 0·1 m.-equiv. of methylene blue added as an indicator, and where n is the *count* determined turbidimetrically but expressed (in millions/ml.) as the equivalent *cell count* of an actively growing standard aerobic glucose culture of the same turbidity. Cultures of *Bact. lactis aerogenes* grown in media containing different carbon sources but with growth limited by the same concentration of ammonium sulphate reach almost identical *turbidimetric counts*, although the *haemacytometer cell counts* may vary several-fold. Hence R as defined above measures the activity of the culture per unit of bacterial nitrogen content and not per cell.

* Imperial Chemical Industries Research Fellow.

In table 1 are given some values for the quantity, R, selected at random from fifty to sixty series of determinations carried out in the space of 18 months. From these figures it is apparent that the variation in R from culture to culture is of the same order as the variations found within a series of determinations made during the growth of one culture. Probably, then, the variations are largely the result of

TABLE 1. THE CONSTANCY OF THE REDUCING POWER IN A GLUCOSE-
AMMONIUM SULPHATE MEDIUM

culture	count (millions/ml.)	R (units as defined) in the text)
1	63	14·5
1	95	13·1
1	150	13·1
1	195	13·2
1	208	13·7
1	243	15·8
1	325	15·5
2	250	13·8
3	220	14·9
4	280	13·5
5	203	14·6
6	202	12·9
7	198	13·2
7	640	13·3
8	62	14·2

TABLE 2. VARIATION OF REDUCING POWER WITH pH IN GROWING CULTURES

series	pH	R
1	7·25	15·0
1	7·2	14·8
1	7·1	15·1
1	7·05	15·5
1	6·9	15·1
1	6·7	16·5
1	6·6	15·8
1	6·4	15·8
2	6·35	13·6
2	6·25	13·4
1	6·2	13·4
2	6·1	12·7
2	5·95	11·5
2	5·8	7·6
2	5·6	4·7

slight random alterations in the experimental conditions and are not due to any systematic changes in the strain. Although the mean generation time of the organism in a standard glucose medium may vary between about 28 and 34 min., no correlation with the variation of the reducing power could be found.

Since it is known that the pH of a culture varies appreciably during the later stages of growth, the activity of cells growing at different pH values was measured. The values are given in table 2, which shows that in the pH range 7·3 to 6·3 R is sensibly constant.

All the experimental evidence shows that this enzymatic activity is a rather stable characteristic of the organism. Since it is intimately connected with the rate of oxygen consumption (Lewis & Hinshelwood 1948), the determination of its magnitude under varying conditions is of general interest in the study of bacterial metabolism. Accordingly, the reducing power shown by *Bact. lactis aerogenes* in two sets of circumstances has been investigated: (1) during growth in different media and (2) under different conditions in a given medium. The first part of the work deals with changes in the carbon source in a standard aerobic ammonium sulphate-potassium phosphate medium.

The following simple carbon sources have been studied: glucose, acetate, citrate, fumarate, α-ketoglutarate, L-glutamate, DL-lactate, DL-malate, pyruvate and succinate, and the following compound carbon sources: glucose + certain amino-acids, glucose + DL-malate, glucose + DL-lactate and glucose + acetate. (In addition, the effect of combining most of the other carbon sources with glucose was investigated; but, as will be shown, the results did not warrant further attention.)

THE REDUCING POWER OF CULTURES AT VARYING STAGES OF ADAPTATION

Cells adapted to grow in a glucose medium require readjustment before they can support growth at the optimum rate or exercise their maximum reducing power in almost any different carbon source. This process, however, is rapid in the majority of cases, and it has been customary to say that no training is required when the adjustment is completed within 1 or 2 hr. Nevertheless, considerable changes do occur in the ability of *Bact. lactis aerogenes* to utilize such carbon sources even during this period. Moreover, a suitable nitrogen source is necessary for these improvements to take place, so that they are presumably due to the synthesis either of further enzyme material or of more coenzymes. Figure 1 illustrates the behaviour that occurs in the malate medium.

The transitory values of R are of significance only in so far as they indicate the progress of the readjustments occurring within the cell. Characteristics so easily acquired are often as easily lost, as was found with a malate strain which required readjustment after one passage through a glucose medium. On the other hand, profound and relatively stable changes occur to the cells during the course of training to substrates such as acetate, D-arabinose, fumarate, glycerol, α-keto-glutarate and succinate. Upwards of twenty serial subcultures, or 160 cell divisions, may be required before the growth rate and the reducing power reach their optimum values (Jackson & Hinshelwood 1948; Baskett & Hinshelwood, 1950; and unpublished results).

In figure 2 the percentage development of the growth rate is plotted against the percentage development of R during the course of adaptation to these substances. Although training, as conventionally understood, does not occur in a citrate medium, growth is sometimes abnormally slow, as though the strain were partially untrained. A point has therefore been included in the figure to show the value corresponding to very slow growth in the citrate medium.

As figure 2 shows, during the course of training to a new substrate an increase in the growth rate is always accompanied by a roughly proportional improvement in R. When this reaches its maximum, the growth rate can increase no more.
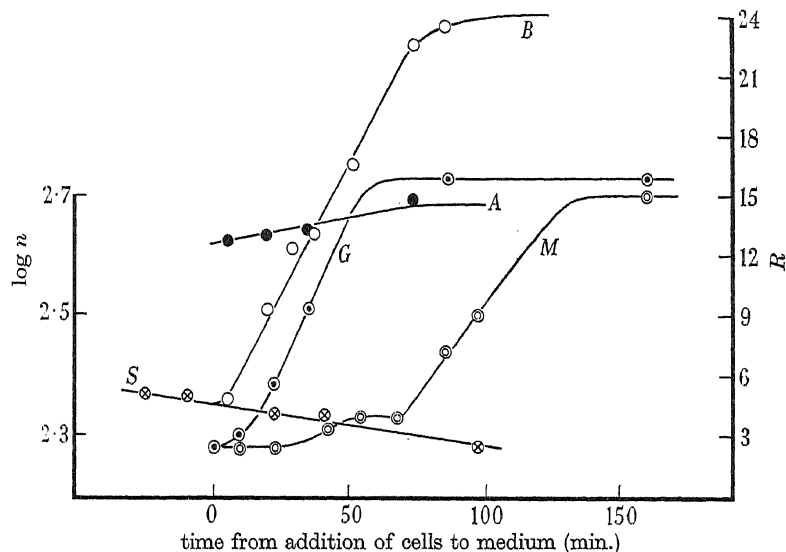
FIGURE 1. Contrasting the adjustment of cells to malate with their immediate utilization of glucose. G and M, growth curves in the glucose medium and in the malate medium respectively. A, reducing power of the glucose culture (showing the standard value for glucose). B, reducing power of the malate culture (rising to the standard value for this culture medium). S, reducing power of the parent suspension in the absence of a substrate.
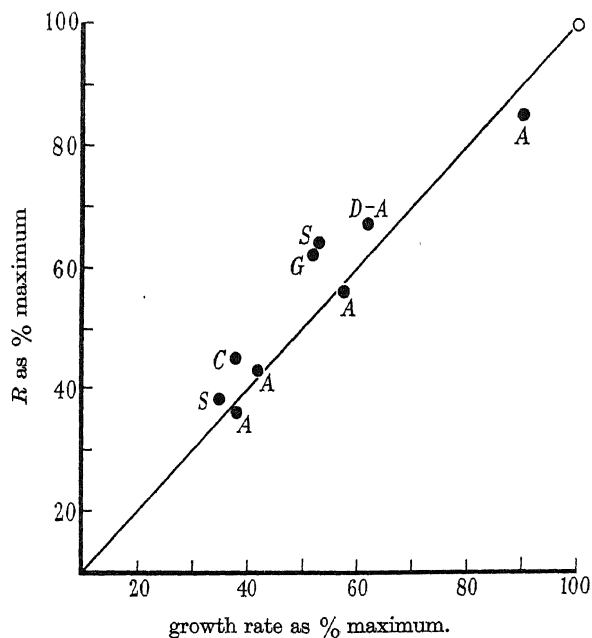
FIGURE 2. The parallel development of R and the growth rate during the course of training to various substrates. Values for cultures containing (A) acetate, (C) citrate, $(D-A)$ D-arabinose, (G) glycerol, (S) succinate.

THE REDUCING POWER OF CELLS IN THE PRESENCE OF DIFFERENT CARBON SOURCES

When all possible adaptation has been completed and the cells are in dynamic equilibrium with any one of the different carbon sources, the value of R for the culture lies in general between 12·5 and 15 units. This relative constancy contrasts with the variations in the mean generation time, which occur between the limits of 90 min. in an acetate medium and 23 min. in a glucose medium enriched with certain amino-acids.

In individual cases this standard value for the reducing power may be obscured by special effects sometimes found in the higher concentration range of the substrate. In view of this, the various carbon sources will be divided into three groups as in table 3.

TABLE 3. CLASSIFICATION OF SUBSTRATES

group 1	group 2	group 3
D-arabinose	DL-lactate	acetate
citrate	DL-malate	
fumarate		
glucose		
L-glutamate		
glycerol		
α-ketoglutarate		
pyruvate		
succinate		
...	...	...
glucose + amino-acids 'A'	glucose + DL-lactate	glucose + acetate
glucose + amino-acids 'C'	glucose + DL-malate	glucose + amino-acids 'B'

Note. Amino-acids 'B' is the combination of twenty amino-acids necessary for growth at the optimum rate in a synthetic medium (Stephens & Hinshelwood 1949). Amino-acids 'A' is the same combination less the sulphur-containing acids, cysteine and cystine. Amino-acids 'C' is a suboptimal combination, also free from the sulphur-containing acids.

The fully trained strains growing on the substrates in group 1 have in each case a mean value of R lying between 12·5 and 15 units. Changes in concentration of the carbon sources do not affect the reducing power to any appreciable extent. For instance, the glucose concentration has been varied between 64 and 16,000 mg./l., with values respectively of 15·2 and 15·5 for R. The growth rate and other kinetic properties of the cell are similarly independent of the glucose concentration in this range, as has been found by many authors using similar organisms (Monod 1942). Not all the substrates under consideration, however, have been investigated over so wide a concentration range as this.

Mean values of R for cultures growing on these substrates are given, together with the mean generation times, in table 4.

The behaviour of the cells in the citrate medium is very complex (Baskett & Hinshelwood 1950), but there does not appear to be any correlation between the slight variations found in R and the normal variation of the mean generation time between 38 and 60 min. The substrate, glucose with added amino-acids 'A', contains the combination of amino-acids, except cystine and cysteine, found to give the optimum growth rate in aerobic synthetic media (Stephens & Hinshelwood

1949). Glucose + amino-acids 'C' is a substrate containing a suboptimal combination of amino-acids. If glucose is added to cultures containing other simple substrates of this class there is little effect on R, and the growth characteristic rapidly become indistinguishable from those in a standard glucose culture.

TABLE 4. THE VALUE OF R FOR CULTURES CONTAINING THE SUBSTRATES
OF GROUP 1

substrate	mean value of R	mean generation time (min.)
glucose + amino-acids 'A'	14·0	23
glucose + amino-acids 'C'	14·0	29
glucose	14·0	32
glycerol	15·0	35
pyruvate	14·5	42
fumarate	15·0	42
succinate	13·0	46
L-glutamate	13·0	48
D-arabinose	14·0	52
α-ketoglutarate	13·5	70
citrate	13·6	38 ⎫
citrate	14·0	41 ⎪ variable
citrate	14·0	46 ⎪ according to
citrate	14·5	49 ⎬ circumstances
citrate	15·0	51 ⎪ of growth
citrate	13·3	55 ⎪
citrate	14·0	58 ⎭

The substrates in group 2 support cultures with rather variable values of R which are generally greater than the standard value of 13 to 15 units. At first sight the behaviour of these substances appears inconsistent with that of the majority of the carbon sources. Further investigation, however, established that the reducing power of these cultures is a function of the concentration of the carbon sources. Figure 3 shows values of R for cultures containing varying concentrations of DL-malate with and without the addition of glucose to the medium. In the series (a) a suspension of cells is added to the complete medium a few minutes before the tests are carried out. (With a suspension small deviations from the normal value of R in growing cultures are to be expected, so that relative values alone are of significance.) In the series (b) the values are for cultures growing logarithmically in media containing the various initial concentrations of malic acid. As the concentration of malate in a culture falls, R approaches more and more closely to the standard value for a glucose culture, although the mean generation time of the cells growing in the malate medium remains sensibly constant. Values for the mean generation time are given in table 5.

It was thought necessary to eliminate the possibility that the sample of malic acid used might contain some growth factor or trace element which would stimulate the oxygen uptake of the cell. Three samples of malic acid of differing purity were investigated, the behaviour of cultures towards each of them being found to be identical (see table 6).

Since centrifuged, washed cells grown in the malate medium show the normal value of R when resuspended in glucose-phosphate mixtures, growth in the malate medium cannot be supposed to lead to a general expansion of the respiratory systems of the bacteria. Further, unadjusted cells suspended in a glucose-malate medium show an enhanced value of R within the time required for the first test,
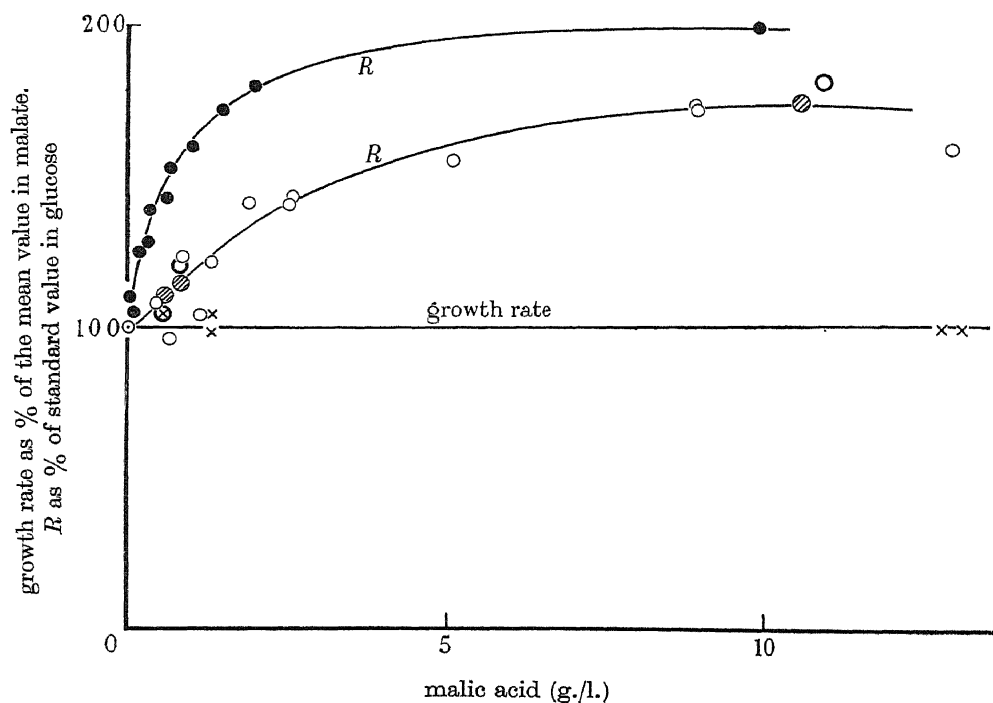
FIGURE 3. Concentration dependence of R in the malate medium. (a) Small filled circles: cells suspended in malate medium shortly before testing; small open circles: cells suspended in glucose-malate medium shortly before testing. (b) Large open circles: mean value of R for cultures growing in malate medium; shaded circles: mean value of R for cultures growing in the glucose-malate medium. Crosses: growth rate in malate medium.

TABLE 5. MEAN GENERATION TIMES IN MALATE CULTURES

g. glucose/l.	g. malic acid/l.	mean generation time (min.)
0	13·0	40·5
0	13·0	40·5
0	1·3	41·0
0	1·3	41·0
0	1·3	38·5
0	0·65	38·5
...	...	...
20	9·0	39
20	2·6	36
20	1·93	32
20	1·3	33
20	0·84	32
20	0·45	32
20	0·00	31

although when suspended in a medium containing malate alone they show a very low value of *R*. Nor does the presence of glucose appreciably affect the adjustment of the cells towards malate in the short time involved. Table 7 illustrates these points.

TABLE 6. A COMPARISON OF DIFFERENT SAMPLES OF DL-MALIC ACID

time from preparation of cell suspension (min.)	*R*			
	sample 1 (10 g./l.)	sample 2 (10 g./l.)	sample 2 further recrystallized (10 g./l.)	glucose (20 g./l.)
3	22·2	20·2	21·0	12·7
15	22·2	22·2	22·2	12·7
38	23·2	22·2	24·6	12·8
66	22·2	22·2	23·2	12·5

TABLE 7. ENHANCEMENT OF REDUCING POWER IN MALATE-GLUCOSE
MIXTURES BEFORE ADJUSTMENT TO MALATE

cells suspended in ...	(1) glucose medium	(2) glucose + malate	(3) malate medium	(4) buffer alone
reduction time (min.)	10·8	6·5	25	40
	10·7	6·8	24	40
R as percentage of value in glucose	100	161	44	26

In the lactate medium the concentration dependence is similar to that in the malate medium. This is shown in table 8.

TABLE 8. THE VARIATION OF REDUCING POWER WITH CONCENTRATION
IN LACTATE AND IN GLUCOSE MEDIUM

mg. lactic acid/l.	mg. glucose/l.	*R*
16,000	0	23
1,430	0	20·5
410	0	19·6
116	0	17·9
33	0	17·0
23	0	14·5
16	0	15·3
13·6	0	14·1
9·5	0	7·7*
9·1	0	6·3*
0	20,000	15·0
...	...	...
0	16,000	15·5
0	64	15·2

* In these cases there is insufficient lactate to complete the reduction of the oxygen and methylene blue, but the cells themselves complete the reduction slowly even in the absence of a substrate.

It may be significant that the maximum value of R in media containing either malic acid or lactic acid tends to a limit roughly twice the standard value.

(Incidentally it should be mentioned that even in the most concentrated malate medium the oxygen concentration at the start of the reduction test was shown directly to be well over 90 % of the standard amount.)

TABLE 9. THE VALUE OF R FOR ADJUSTED CULTURES CONTAINING DILUTE SUBSTRATES OF GROUP 2

substrate	R	mean generation time (min.)
lactate	14–15	56
malate	14–15	42
glucose + malate	14–15	32

TABLE 10. INHIBITION BY ACETATE

culture	g. acetic acid/l.	R	mean generation time (min.)
acetate alone	20	10·5	100
glucose + acetate (untrained cells)	20	6·7	—
acetate alone	1·67	12·5	90
acetate (glucose added in test alone)	1·67	12·3	—

The reducing power of cultures in these media may now with justification be related to two distinct processes. The first is a fundamental respiratory mechanism which contributes to R the standard value of 14 to 15 units. Like the growth rate this only varies in the very lowest concentration range. Both, moreover, depend upon the useful assimilation of the substrate, a process which often can only proceed at the optimum rate after the cells have become adjusted to a new medium. The second process consists of an enhancement up to a twofold limit of the fundamental respiratory process, on whatever substrate this is based. This secondary effect is easily distinguished because it comes into play only at relatively high concentrations of certain very specific substrates such as malate or lactate. The increased rate of oxygen consumption in these high concentration ranges is not paralleled by any improvement in the rate of growth (see table 5). In point of fact, this increased activity in glucose-malate cultures is quite uneconomic, the growth rate dropping with increasing malate concentration, that is, with increasing oxygen consumption.

To sum up, in cultures containing these apparently anomalous carbon sources at moderately high dilutions, R has the standard value of approximately 14 to 15 units. That part of the oxygen uptake represented by the higher values of R in more concentrated media appears to have no beneficial effect upon the rate of synthesis of cell material. Thus, for correlation with rates of growth, the relevant value of R in these culture media may be taken as the 14 to 15 units found at the higher dilutions. These values are given in table 9.

The carbon sources in group 3 also appear at first to possess anomalous properties in that they inhibit certain cell processes. Fully trained cultures in a concentrated acetate medium (20 g./l. acetic acid) show a value of R of only 10·5 units.

The acetate appears to inhibit the reduction of the methylene blue in the test, the fading of the dyestuff being visibly prolonged. In a dilute acetate medium, however, the growth rate is slightly more rapid and the value of R rises to about 12·5 units, which is maintained when glucose is added to the test (see table 10).

In a culture containing the mixture glucose and amino-acids 'B', the sulphur-containing amino-acids, cysteine and cystine which are present, so inhibit the reduction of the dyestuff that it is difficult to obtain reproducible reduction times. If these two amino-acids are excluded normal values of R are observed.

DISCUSSION

When due allowance is made for the special effects associated with the carbon sources of groups 2 and 3, there appears to be for fully adapted cultures a basic mean value of R which varies only between 12·5 and 15 units, although the mean generation times range from 23 to 90 min. (see figure 4). On the other hand, during the course of adaptation to any given substrate R varies in parallel with the growth rate of the culture (see figure 2).
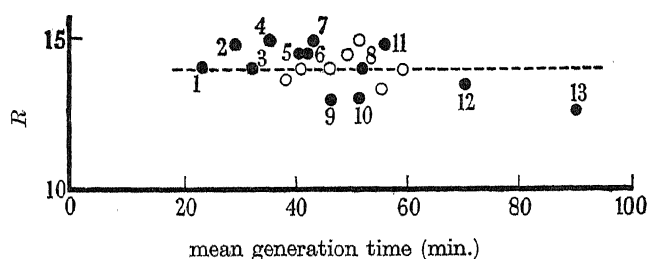
FIGURE 4. The constancy of R in fully-adapted aerobic cultures at 40° C. Carbon sources: (1) glucose-amino-acids 'A', (2) glucose-amino-acids 'C', (3) glucose, (4) glycerol, (5) DL-malate, (6) pyruvate, (7) fumarate, (8) D-arabinose, (9) succinate, (10) L-glutamate, (11) DL-lactate, (12) α-ketoglutarate, (13) acetate. The unfilled circles represent citrate cultures growing under various conditions.

This standard value of R for fully adapted cells implies a constant rate of oxygen uptake per unit bacterial nitrogen content. It suggests that the optimum growth rate attainable by the adaptation of *Bact. lactis aerogenes* to a given carbon source is fixed by a rate-limiting step in the respiratory system of the cell. When the cells are first transferred to a new medium the proportions of various enzymes are unsuitable for the optimum utilization of the substrate, which will involve a specific sequence of reactions. The oxidative mechanisms, therefore, are not fully utilizable. At first the extent to which these proportions are readjusted determines both the growth rate and the degree to which the oxidizing mechanisms are used. During the course of adaptation, in fact, both R and the growth rate will at first be governed by the initial steps in the metabolism of a new substrate, and the two develop in parallel with one another. The conversion of different substrates into cell material, with the synthesis of the vast variety of compounds thereby involved, will obviously require different quantities of oxygen. For instance, more free energy should be

expended in the synthesis of bacteria from acetate than from glucose, and the free energy in aerobic metabolism is derived largely from oxidation. The existence of lactic fermentation shows that less free energy is derivable by oxidation from lactate than from an equivalent quantity of glucose and oxygen. There are good reasons, then, for supposing that, when a constant rate of oxygen uptake is the limiting factor, fully adapted cultures should have widely differing growth rates in different types of medium.

Whether more precise experiments would reveal a rigid constancy of R is doubtful. If it were so, it would imply that the respiratory system of the bacteria always formed a fixed proportion of the total cell mass or of the total bacterial nitrogen content. Since adaptation results in changed proportions of some other enzymes, an absolutely fixed proportionality is in principle unlikely. But during the course of adaptation to a new carbon source only certain enzyme ratios change, namely, those immediately responsible for the metabolism of this substrate. If, on the other hand, there is an actual change in the hydrogen acceptor provided, this might well result in an adaptation affecting the respiratory enzymes themselves. This may be what occurs in the adaptation of the cells to aerobic nitrate metabolism where a value of R of 26 units is found (Lewis & Hinshelwood 1948).

REFERENCES

Baskett, A. C. & Hinshelwood, Sir C. 1950 *Proc. Roy. Soc.* B, **136**, 520.
Jackson, S. & Hinshelwood, Sir C. 1948 *Trans. Faraday Soc.* **44**, 568.
Lewis, P. R. & Hinshelwood, Sir C. 1948 *Proc. Roy. Soc.* B, **135**, 301.
Monod, J. 1942 *La croissance des cultures bactériennes*. Paris: Hermann et Cie.
Quastel, J. H. & Whetham, M. D. 1924 *Biochem. J.* **18**, 519.
Quastel, J. H. & Whetham, M. D. 1925 *Biochem. J.* **19**, 645.
Stephens, D. & Hinshelwood, Sir C. 1949 *J. Chem. Soc.* p. 2516.
Thunberg, T. 1920 *Skand. Arch. Physiol.* **40**, 1.

The hypothalamic control of food intake in rats

By G. C. KENNEDY

National Institute for Medical Research, London, N.W. 7

[Plates 24 and 25]

The effect on feeding behaviour in the rat, of permanent electrolytic lesions in the hypothalamus which cause obesity, has been studied.

Soon after operation, the animals were extremely voracious. As the obesity became established, the voracity was reduced in degree and finally disappeared. Young unoperated rats could not be made to exhibit such voracity even after starvation.

The young unoperated animals showed little discrimination against unattractive diets. When fed on meal mixed with an equal weight of kaolin, they obtained a normal nutrient intake, and maintained their weight by doubling the bulk eaten. In the voracious stage, operated animals adapted in this way more quickly than normal; in the obese stage they adapted more slowly and only after loss of considerable body fat. Older unoperated animals, which were moderately obese, resembled the operated ones in their feeding reactions.

It is suggested that the effect of hypothalamic damage is to emphasize two components of the urge to eat, a primitive voracious one, and a more discriminative component, either of which may be dominant according to the state of the fat stores. The simplicity of the young rats' reactions appears to be due to closer co-ordination of these components by the hypothalamus. Modification of the hypothalamic control of energy intake may be a pre-requisite of obesity in general.

Reasons are advanced for regarding the hypothalamic mechanism as being sensitive to chemical changes in the blood.

INTRODUCTION

Obesity has long been known to be associated in some way with the hypothalamic-pituitary complex. Brobeck (1946) attributes the first clinical description of rapidly increasing weight associated with a tumour of this region to Mohr in 1840. Babinski (1900) and Fröhlich (1901) described obesity in adolescents associated with destructive pituitary lesions, and attributed it to diminished secretion of the gland. However, in 1904, Erdheim expressed the dissident view that the source of the metabolic disturbance in Fröhlich's syndrome lay in the hypothalamus, and the aetiology has remained a subject of controversy ever since. The recent experimental studies of adiposity following exclusively hypothalamic damage in animals, notably by Ranson, Brobeck and their co-workers, have done much to simplify the position.

The experimental approach to the 'hypothalamic' theory of obesity may be said to have begun with the work of Camus & Roussy (1913, 1920, 1922) on the dog, and of Smith (1927, 1930) on the rat. These workers showed that removal of the pituitary was not followed by increased adiposity, whether the removal was partial or complete, so long as the base of the brain was not damaged. Where obesity did occur, the hypothalamus was found to be affected. Various workers since then have recorded obesity following injury which appeared to be mainly hypothalamic, although, because of the subtemporal routes of operative approach generally used, damage to the gland could not be excluded with certainty.

In 1939, Hetherington & Ranson, using the Horseley-Clarke stereotaxic instru-

ment, were able to produce obesity in the rat by electrolytic lesions, introducing the electrode downwards through the parietal bone to avoid suspicion of damaging the pituitary. Brobeck, Tepperman & Long (1943) confirmed this work, which has subsequently been extended to the cat and the monkey.

The lesions produced by Ranson were fairly widespread in the ventro-posterior hypothalamus, but the most effective ones always appeared to involve the ventro-medial nuclei, or fibres which Ranson believed to leave their ventro-posterior aspect. Hetherington (1943) showed that hypophysectomy either before or after the hypothalamic operation did not prevent the increase in weight, and concluded, '...since neither total nor partial hypophysectomy produces adiposity, or prevents its appearance after hypothalamic damage is done, it is not likely that the hypophysis is involved in the production of obesity often associated with injury to structures in the pituitary region'.

Brobeck, with rather more discrete lesions, obtained extreme obesity by damaging the same area of the hypothalamus. Investigation showed that in most cases the food intake was increased, sometimes to two or three times normal, and at the same time the spontaneous activity of the animal was reduced. Sometimes the hyperphagia, as Brobeck has called it, appeared to be no more than could be expected from the excess weight of the rat, but in other cases the appetite actually decreased as the rat became fatter until a balance was attained.

There is a marked difference in the relative importance attributed to the intake and activity factors by the two schools of workers. Hetherington & Ranson (1942) record that some of their most obese rats did not show an increased food consumption, but that most of the animals in which they damaged the hypothalamus, whether they became obese or not, showed greatly reduced activity. Brobeck's more localized lesions produced obesity of more dramatic onset, and usually of greater degree, and in his animals hyperphagia was often obvious from the time of recovery from the anaesthetic; there can be little doubt that it was the major cause of obesity. In view of the difference in the extent of the lesions, it is possible that Ranson was dealing with multiple effects, and Brobeck with a single one. Investigation of other factors which might be expected to produce economy of energy, such as changes in intestinal absorption or intermediate metabolism, has shown only minor and inconstant abnormalities, which appear quite inadequate to explain the great rates of weight increase which have been observed. These investigations have been reviewed by Brobeck (1946).

To summarize, it appears that exclusively hypothalamic damage can cause extreme obesity which is adequately explained by the increase of food intake observed during its development. In its most characteristic form, this hypothalamic obesity is not accompanied by any other metabolic change, or decrease in activity, which might significantly influence the 'metabolic equation', nor does the pituitary appear to play any part.

The present investigation has been undertaken to examine more closely the nature of the disturbance in feeding behaviour which Brobeck has called hyperphagia, and to find whether any signs of the activity of a hypothalamic 'centre' can be detected in the normal animal.

Methods

Animals and operative procedure

Initially, two different Wistar strain albino rats were used, one of them differentiated in table 1 by the prefix B, and a hooded strain with the prefix H. Litter-mate controls were always used. The strains appeared to vary in their tendency to become fat under normal conditions, but no difference was detected in their capacity to develop obesity after operation. Later work has therefore been restricted to Wistar strain albinos, indicated by plain numbering, to which all normal growth data quoted refer. The operations were carried out under intraperitoneal bromethol anaesthesia, using a Horseley-Clarke stereotaxic instrument in the manner described by Ranson and Brobeck. The only essential post-operative precaution appeared to be the maintenance of body temperature by warming the room to 75 to 80° F.

Objective estimation of the degree of obesity

Assessment of obesity may be difficult where the rate of skeletal growth of an animal differs much from that of its litter-mate control. The relation of weight to nose-anus length was therefore investigated in the normal animal. Ranson found the relation between weight and the cube of length to be inconstant, so we based our analysis on the more general function expressed by Huxley's allometric growth equation $y = bx^{\alpha}$ with weight and length as the variants y and x. All lengths were determined under anaesthesia. Using data from eighty-three animals measured within the same two weeks, analysis for linear regression of log weight on log length gave a correlation coefficient of 0·9868 with $P < 0·01$. The slope of the regression line, which is shown in figure 1, is approximately 3·25. This is consistent with Huxley's equation, expressed in the logarithmic form $\log y = \log b + \alpha \log x$. The slope is sufficiently different from 3 to explain Ranson's difficulty with a simple cubic relation, especially as his published data on control animals yielded a regression line of very similar slope to that of figure 1.

From the length of an obese animal, reference to the curve gives the 'regression weight' of a control of equal length; I have referred to weights calculated for obese animals in this way as 'ideal weights'.

Determination of the fat and water content

Fat and water content of the rats was determined separately on the skin with the adherent subcutaneous tissue, and on as large an aliquot as possible—usually about 50 g.—of the remaining minced carcass. Before mincing, the gut was opened and wiped clean of contents, in order that comparisons with animals on high-residue diets should be valid. The water was first distilled off under reflux with benzene, and collected in an extractor tube arranged to spill back the benzene into the distillation flask. Twelve hours extraction gave a constant volume of water. The dehydrated tissue was then further extracted with the same benzene in a Soxhlet apparatus. Re-extraction with ether was found to give a negligible increase in fat yield. Finally, the residue was dried and weighed.

This method makes it possible to determine fat, water and dried residue fractions directly on the same large sample of tissue. We found the calculated fractions of the carcass as a whole yielded a total which differed little from the body weight directly measured. The mean error in determinations was $2 \cdot 403 \pm 0 \cdot 224 \%$.
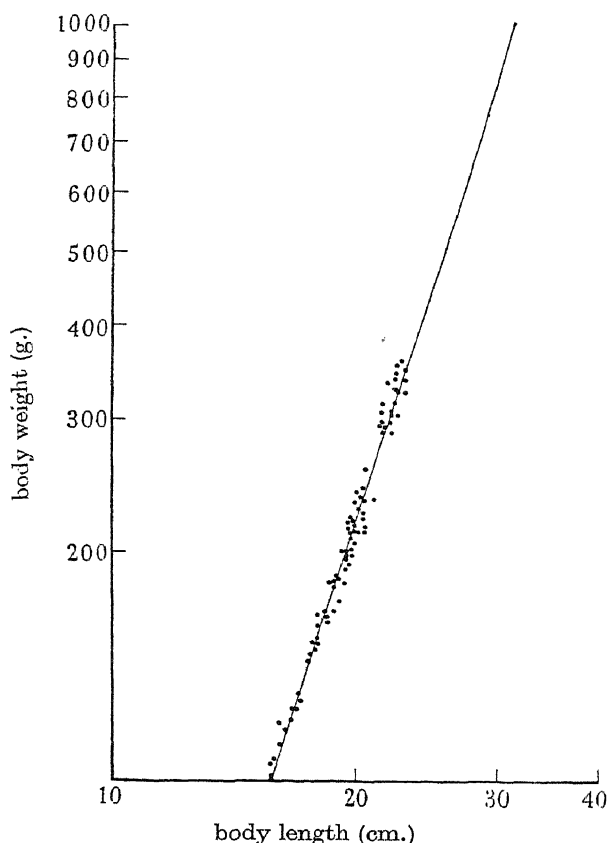
FIGURE 1. Relation between log weight and log length of normal albino rats.
Slope $= 3 \cdot 25$; $r = 0 \cdot 9868$; $P < 0 \cdot 01$.

Feeding of adulterated diets

Adolph (1947) showed that young rats presented with a diet containing varying percentages of an adulterant such as kaolin or cellulose responded within 2 or 3 days by increasing their overall consumption until the amount of nutrient eaten was restored to its original level. His rats were able to make this adjustment until something between 30 and 50 % of diluent had been added to the diet, after which the mechanism apparently failed; he suggested that the limiting factor was gut size.

This technique has been applied to young normal rats of different age groups, and to hyperphagic animals, and marked variations in feeding reactions have been observed. The test diets employed were the stock meal used in this laboratory, powdered finely to prevent selection, and mixtures of this with kaolin containing respectively 75, 50 and 25 % of stock diet. Rats were kept on wire grids over clean
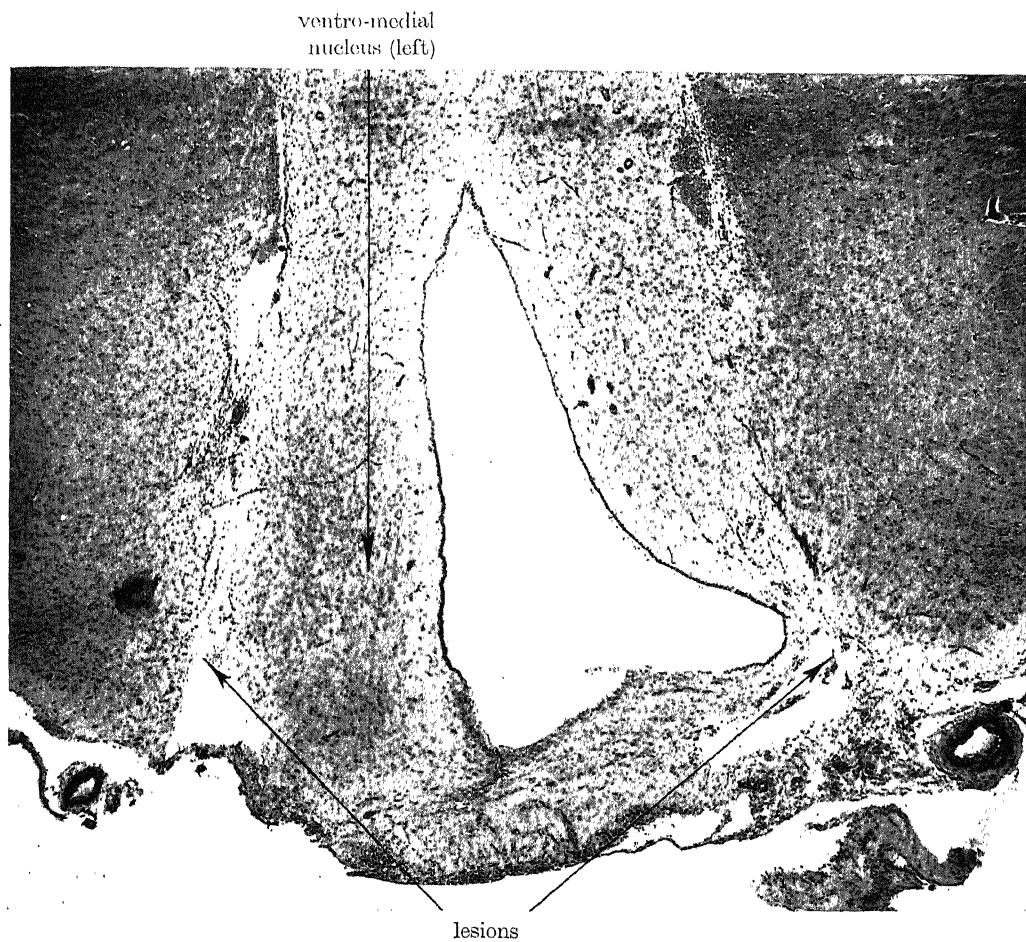
FIGURE 2. Lesions in hypothalamus of rat 230, 19 weeks after operation (magn. × 50). Section through tuberal region, showing destruction of ventro-medial nucleus and distortion of third ventricle on right side, with little damage to the nucleus on the left.
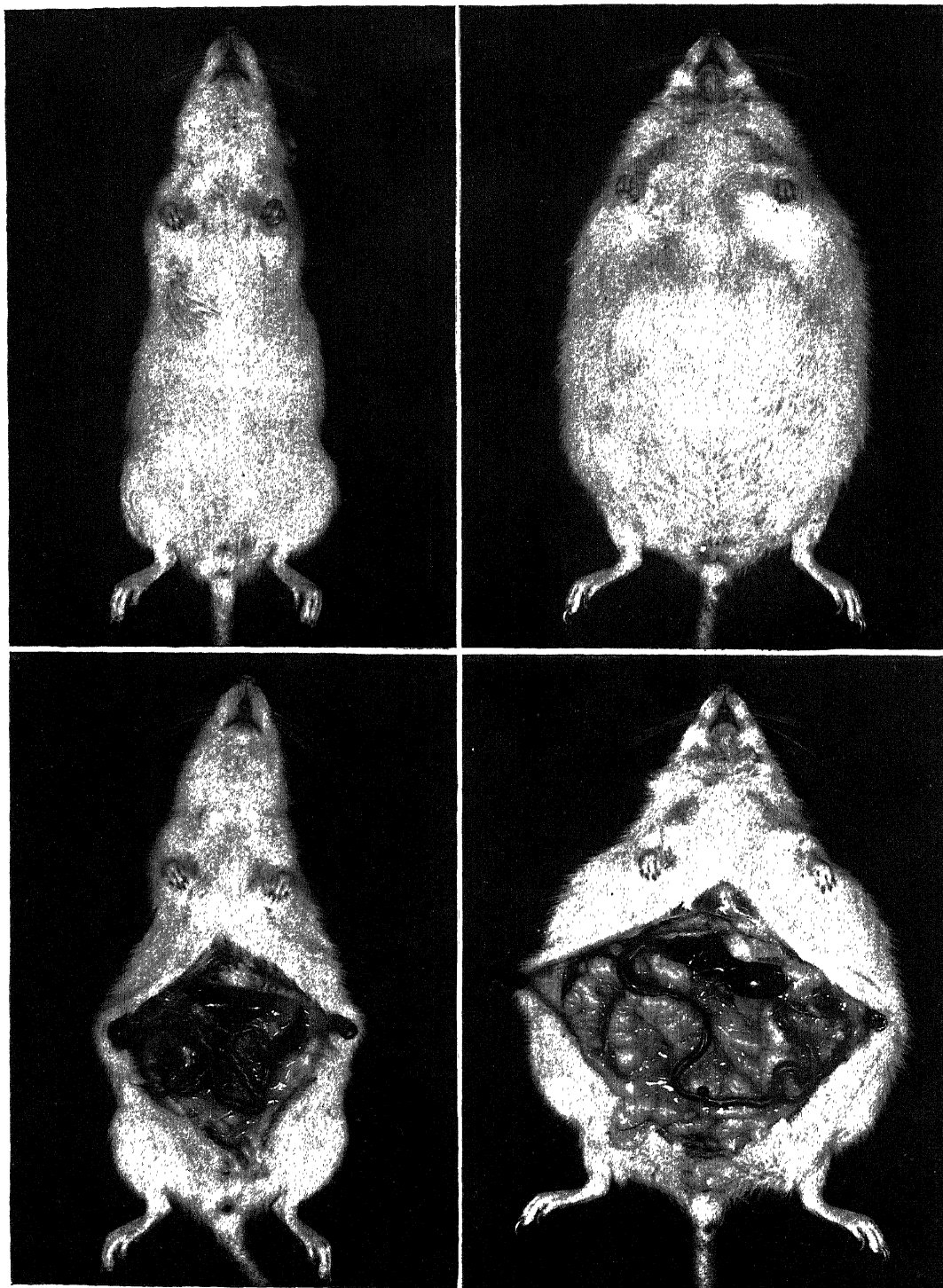
FIGURE 3. Comparison of obese rat 230 and control 224 at autopsy, 19 weeks after operation.

paper to detect spilling, most of which was prevented by placing the food cups inside deep metal containers. All spilt food was collected and weighed. Room temperature was kept between 75 and 80° F.

RESULTS

Position of effective lesions

During the initial stages of the work, in a series of 120 animals, bilaterally symmetrical lesions were placed in as many positions as possible in the hypothalamus and rostral part of the mid-brain. Although these animals were kept for 6 months, none of them showed any sign of becoming obese, except where the lesion was in the area described by Ranson and by Brobeck, in the lateral part of the tuberal region. It has been my experience also that the most effective lesions lie ventro-lateral to the ventro-medial nuclei, and often appear to cause little direct damage to the nuclei themselves. It has been impossible to confirm the finding of Hetherington & Ranson, that obesity could be produced by lesions caudal to these nuclei, even when two large lesions were placed on each side of the mid-line. Figure 2, plate 24, shows the lesions in rat 230, and figure 3, plate 25, the degree of obesity exhibited by the animal compared with its control at autopsy, 19 weeks after operation, illustrating the enormous increase in depot fat seen in these animals.

Analysis of carcasses

The water, fat and dry-residue percentages of ten obese rats and their controls are given in table 1. Comparison is helped by expressing the percentages in terms of the ideal weight, when it is seen that the body water is unaffected by the operation. The dry residue is slightly increased, more markedly in the two fattest animals 130 and 267, but the increase is of a much smaller order than that in neutral fat, and may well be explained by changes in musculature secondary to the greatly increased weight carried. Comparison of the 'ideal weights' with the weights of the controls makes it clear that there is no consistent change in skeletal length, such as might be expected if there were any pituitary disturbance.

In table 2 a comparison is made of three possible methods of estimation of adiposity. The regression method, in which the ideal weight is subtracted from the observed weight, gives a better agreement with the results of fat analysis than direct comparison with the litter-mate control. This is especially marked in the case of animal 173, which was much shorter than its control, and although moderately obese, was almost equal to the control in weight.

The increase of weight in these animals is fully accounted for by excessive fat deposits, and apparently neither water retention nor 'growth' in its usually accepted sense plays any part.

Feeding behaviour on stock diet

Animals with hypothalamic lesions often eat voraciously immediately after recovery from the anaesthetic. Such behaviour was noted by Brobeck, who recorded that some of his animals ate more in the first 24 hr. after operation than in any subsequent day. While this was also our experience, we noted that such voracity

was not a reliable prognostic sign of future obesity, as the increased appetite was often not maintained for more than a few days. On the other hand, the permanent type of hyperphagia was usually recognizable by the end of the first week, when an increase in weight of 50 g. had often occurred, and the daily food intake was from

TABLE 1. ANALYSIS OF OBESE RATS AND THEIR CONTROLS

rat no.	weight (g.)	ideal weight (g.)	percentage water		percentage fat		percentage residue	
			actual	ideal	actual	ideal	actual	ideal
126o	397	315	52·2	64·5	30·6	38·6	19·6	24·7
130o	516	290	33·4	59·3	48·5	86·5	16·1	28·6
129c	318	325	64·4	63·0	12·0	11·7	24·2	23·7
165o	304	224	46·8	63·6	35·4	48·0	19·8	26·9
169c	215	224	64·6	62·0	14·2	13·7	24·0	23·0
173o	376	335	51·25	57·6	30·8	34·6	19·2	21·5
171c	372	390	57·6	55	17·2	16·3	20·5	19·5
176o	292	224	47·2	61·6	42·0	54·7	18·3	23·8
177c	216	224	65·5	63·1	17·5	16·9	23·9	23·0
230o	356	185	33·3	64·0	53·5	100·3	13·5	26·0
224c	203	210	61·1	59·0	18·2	17·5	22·5	21·8
267o	524	285	37	68·1	46·4	85·1	18·4	33·7
266o	336	280	52·6	62·9	27·0	32·3	21·2	25·4
268c	254	270	63·2	60·0	13·0	12·25	27·0	25·4
B033o	356	224	40·8	64·7	44·3	70·5	16·5	26·2
B034c	186	195	63	60·0	13·5	12·8	24·5	23·3
H013o	340	—	44·8	—	40·5	—	13·9	—
H015c	210	—	58·5	—	19·8	—	22·3	—

o = obese; c = control.

TABLE 2. METHODS OF ESTIMATION OF ADIPOSITY OF OPERATED RATS

rat no.	weight minus weight of control	weight minus 'ideal weight'	fat content of operated minus control
126	79	82	83·5
130	198	226	213
165	89	80	77
173	4	41	52
176	76	68	85
230	153	171	154
266	82	56	57
267	270	239	230
B033	170	132	133

twice to three times that of the control animal. This represented in most cases a maximum level of intake, which was maintained, often with little alteration, for several weeks and then gradually declined, until eventually little difference could be detected between the intake of the obese and control animals. When the food intake returned to normal levels, the degree of obesity was apparently maximal, and any growth observed merely reflected the normal increase shown by the control animal. These points are illustrated in figure 4. The hyperphagic period in animal 184 lasted about 60 days, although it was temporarily interrupted at 30 days. The temporary

fall in intake was probably due to slight change in environmental conditions to which these animals are very sensitive. It coincided with a rather less marked reduction of intake by the control animal. Although a few of our very obese animals never returned to a level of food intake as low as that of their controls, we repeatedly observed that increases of the order of 50 % over control weight were maintained
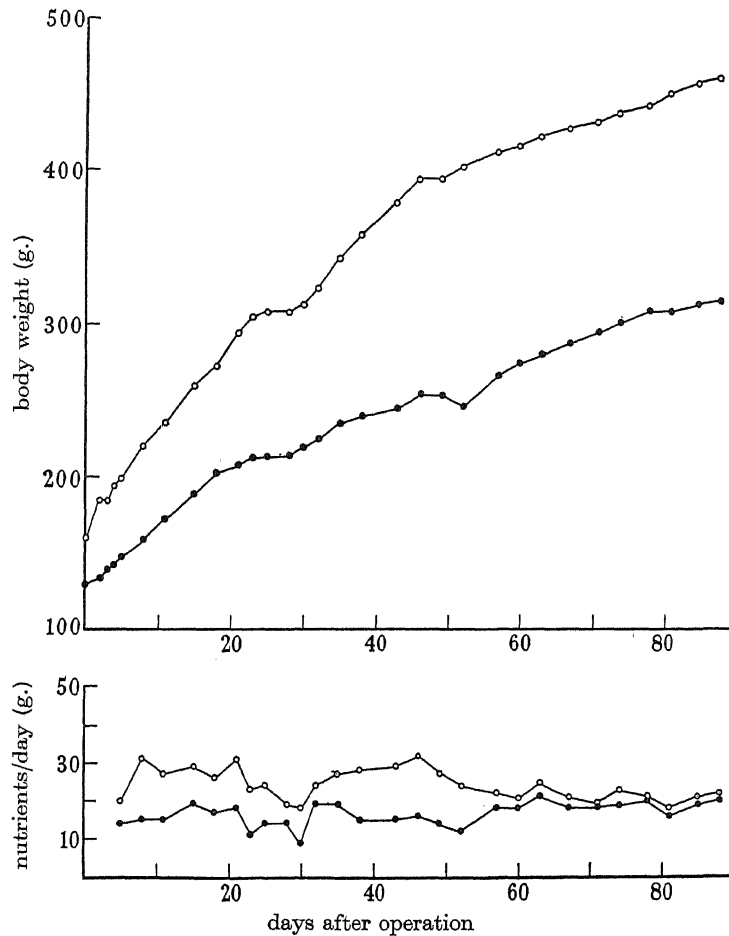
FIGURE 4. Growth and food intake of operated rat 184 and control rat 183 for 90 days from operation. o 184; ● 183.

apparently indefinitely on a hardly detectable excess of food intake. At the same time it was clear that this represented an equilibrium dependent upon the presence of excessive fat depots, and not a permanent recovery from the hyperphagic disturbance, as any reduction of weight by restricted diet resulted in an immediate return of the hyperphagia when unrestricted feeding was allowed.

Reactions to adulterated diets

In a preliminary experiment we tested Adolph's findings, keeping as exactly as possible to his criteria for animal weight and period of test. Four groups each of five young male albino rats between 150 and 300 g. were given in turn either powdered

stock diet, or the same diet diluted with 25, 50 or 75 % of kaolin by weight. The test periods were each of 7 days, food intake being measured only during the last 4 days, and a week was allowed on normal diet between each test. The allocation of groups to diets was in accordance with a Latin square. Figure 5 shows the mean daily intake of all four groups expressed as total bulk eaten, and as calculated nutrient intake, on the assumption that intestinal absorption is unimpaired. (Separate tests have shown that even after 4 weeks on a 75 % kaolin diet, the rat does not show any significant failure of absorption.)
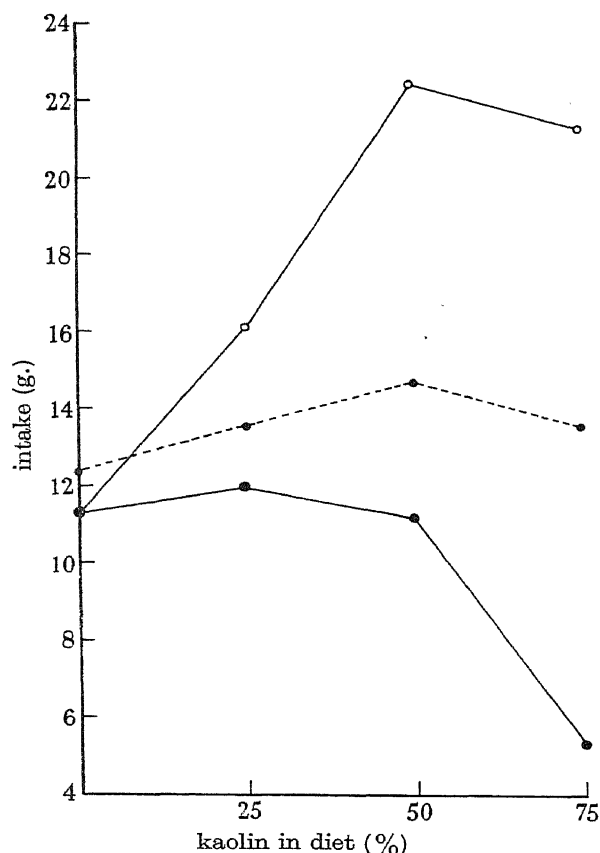
FIGURE 5. Effect of dilution of diet with kaolin on daily intake of normal male rats. o — o Total solids eaten; •—• nutrient intake; • – – • water intake.

Analysis of variance showed no significant change in nutrient intake to 50 % adulteration, but a highly significant fall in intake at 75 % of kaolin, thus confirming Adolph's finding, at least on young animals.

The difference in reaction between obese and normal animals to adulterated diets was most typically shown with 50 % kaolin mixtures, and when the obesity was fully developed. In figure 6 it will be seen that the control animal (183) restored its nutrient intake within about a week, and maintained a steady weight little different from its former level. About 4 weeks was necessary before the nutrient intake of the obese animal equalled that of the control. At the end of this time the weight

became steady, but the obesity was considerably decreased. Further reduction of weight by restricting the diluted diet was followed on derestricting by restoration of this steady weight level, which appeared to be characteristic for the individual and the particular diet.
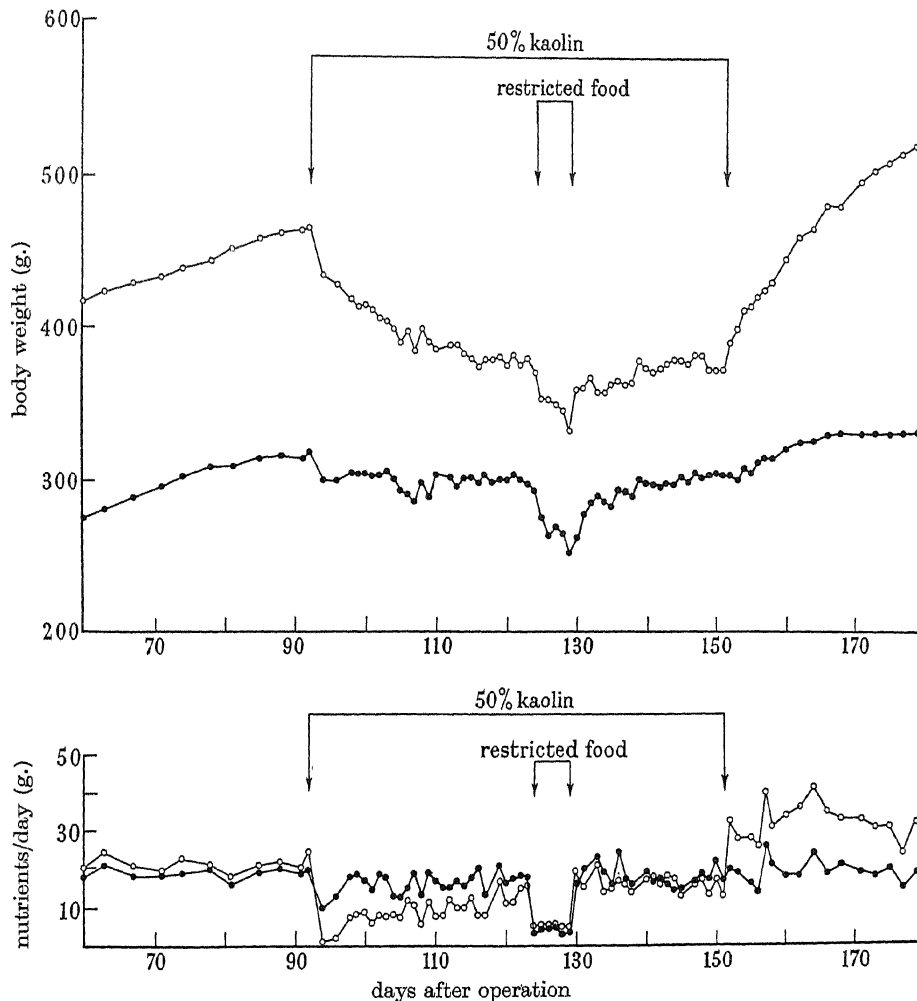
FIGURE 6. Reaction of animals 184 (operated) and 183 (control) to 50% kaolin mixture. o 184; ● 183.

If the kaolin diet was given during the initial hyperphagia, before the obesity was fully developed, the reaction differed markedly. The adjustment of intake was much more rapid, and little fat appeared to be lost. In figure 7, animal 267 lost about 50 g. in apparent body weight on the first 2 days, during which it ate no food. As it had until then been forcing 40 g. of food a day into its distended gut, this probably merely reflected gut emptying. Little more weight was lost, and the food intake, as in the previous case, remained identical with the control until unrestricted feeding of undiluted food was restored, when the hyperphagia reappeared in its original intensity.
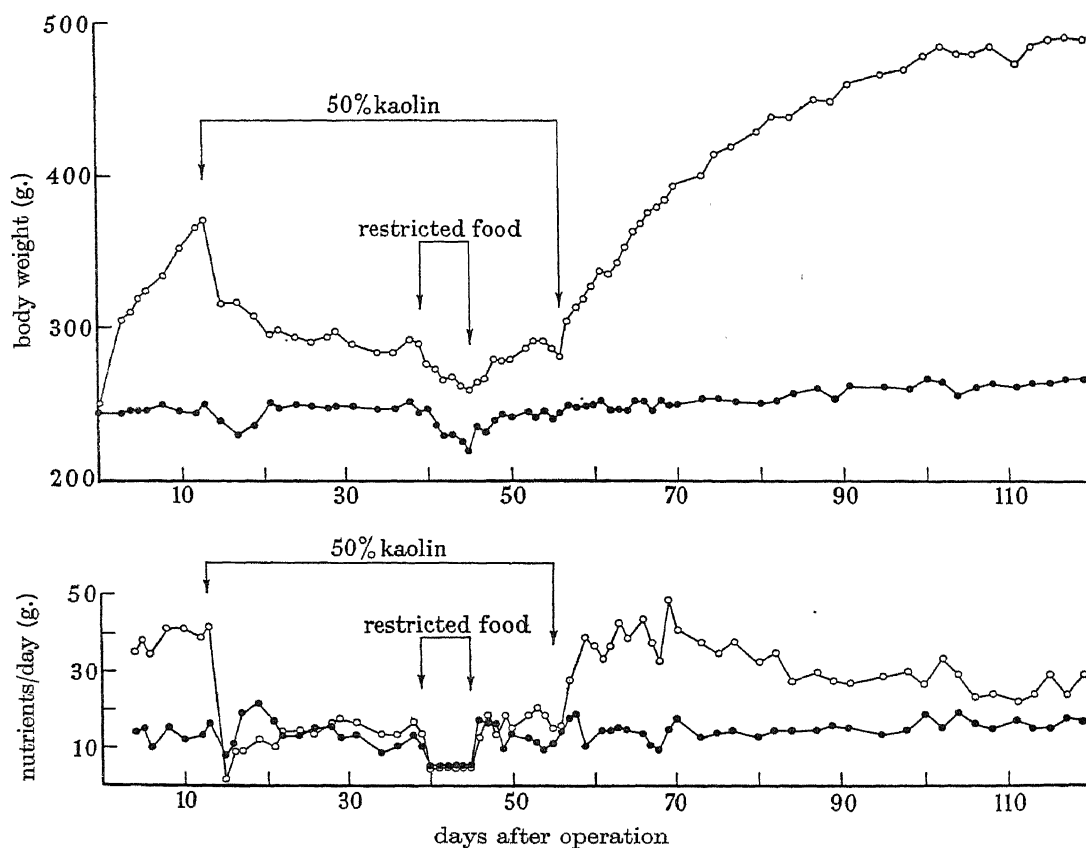
FIGURE 7. Reaction of animals 267 (operated) and 268 (control) to 50 % kaolin diet introduced during hyperphagia. o 267; • 268.

Reactions of unoperated animals to adulterated diets

The observation that the older, rather fatter animals tended to react to test diets in a manner more like the operated rats than like the young ones originally studied, suggested a reinvestigation of the behaviour of 'naturally' fat rats. 102 male rats of weights varying from 150 to 450 g. were given the 50 % kaolin mixture for 14 days, by which time they all appeared to have reached a steady weight. In figure 8 the weight lost on the kaolin mixture is plotted against the initial weight, of which it is clearly a function. It will be seen that animals below 300 g. can, on the whole, maintain their weight on 50 % kaolin; when the diet was diluted to 75 % kaolin, only young growing rats below 200 g. could tolerate this without weight loss.

Adolph, who used young rats exclusively, found that deprivation of food either compulsorily or voluntarily as a result of refusal of adulterated diets, was not followed by increased intake on return to unrestricted feeding. He concluded that 'previous privation of food is little or no stimulus to eating'. Again, our experience has been different with young and old animals. Figure 9 shows the mean daily food intake of two groups of five male rats during the last of their 4 weeks on 75% kaolin, and after their return to normal stock diet. The rats of initial weight of the

order of 200 g. showed little increase in intake during recovery, but the older rats showed clear hyperphagia. Analysis of variance shows the difference 'between weeks' in the heavier animals to be significant. In this experiment the younger animals had lost little or no weight during the kaolin feeding. Moreover, when the
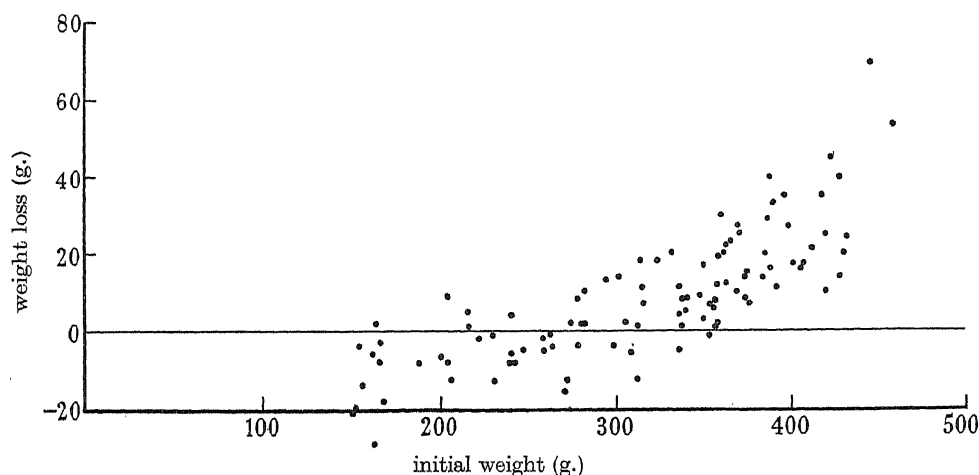
FIGURE 8. Weight loss of unoperated male rats fed for 14 days on 50 % kaolin diet. Animals which continued to grow are indicated by a negative weight loss.
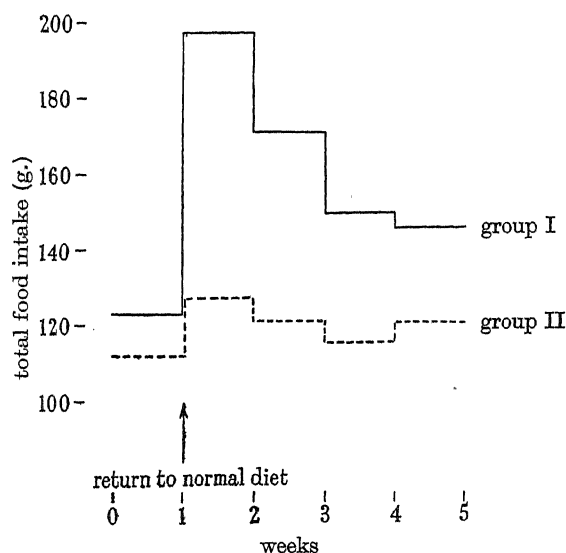
FIGURE 9. Mean weekly food intake of groups of 5 unoperated male rats on return to stock diet, after 4 weeks feeding on a 75 % kaolin mixture. Group I, mean initial body weight 372 g. Group II, mean initial body weight 227 g.

deprivation of food was compulsory, there was still no sign of hyperphagia in the young animals on return to unrestricted feeding. In figure 10 the daily food intake of three groups of five animals of mean weight approximately 400, 350 and 200 g. are compared with similar control groups recorded at the same time. The experimental groups were restricted to a daily intake of 5 g. of food for 3 days, then 10 g.

G. C. Kennedy
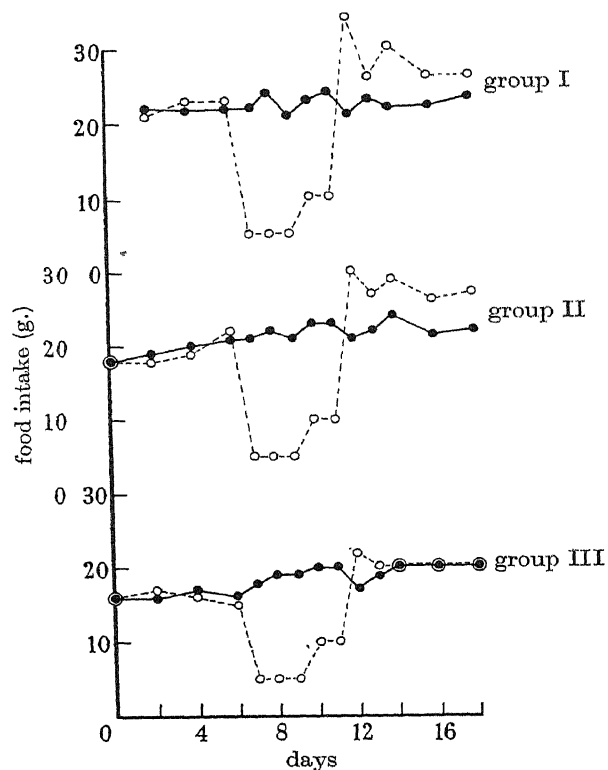
FIGURE 10. Increased food intake in normal rats on unrestricted feeding following partial starvation. Groups of five males with control groups of similar weight. Group I, initial body weight 382 to 421 g. Group II, initial body weight 332 to 361 g. Group III, initial body weight 184 to 207 g.
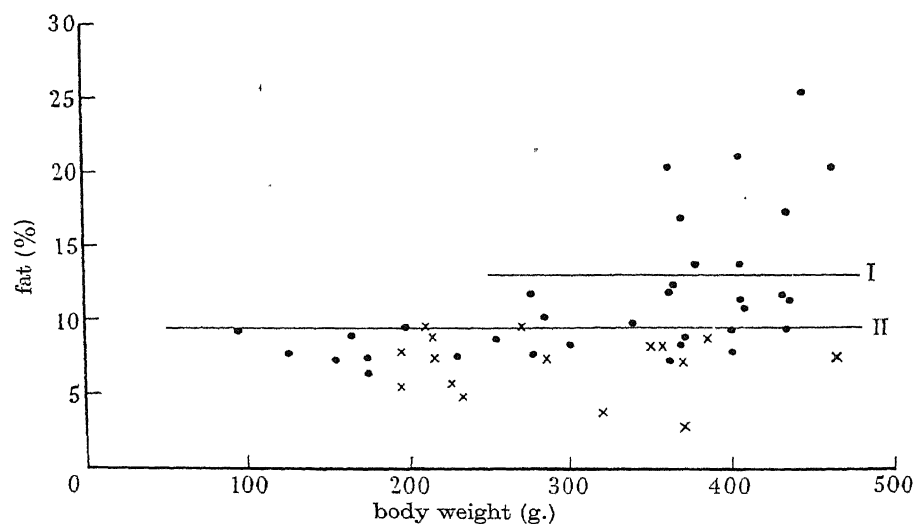
FIGURE 11. Reduction in percentage body fat of unoperated rats maintaining a steady weight on adulterated diets. ● Control rats on stock diet. × Rats fed on the 75 % kaolin mixture. The final fat percentage is plotted against the initial weight on stock diet, for comparison with the controls. I, II, maximum recorded fat percentage in animals eating 50 % and 75 % kaolin diets.

for a further 2 days. On return to unrestricted feeding, the heavier groups showed hyperphagia, the lightest group none.

It was therefore decided to investigate the normal fat content of our colony rats at different weights, and figure 11 shows the result of thirty-four fat analyses, plotted against body weight. The range of percentage fat found increases markedly with increasing weight and age, although as the precise age of the animals was unknown an exact statement of the relationship with age is impossible. The stock was young and growing, however, and in general, weight was a function of age. On the same diagram the percentage body fat found in animals at the end of the two previous experiments with kaolin diets is plotted against the weight at the beginning of the experiment. It will be seen that there is a marked reduction in the range of body fats recorded in rats above 300 g. fed on 50 % kaolin diet, while all the rats fed on 75 % kaolin mixture lost fat down to the level characteristic of 200 g. animals.

DISCUSSION

Brobeck's work made it clear that the obesity in his rats was due to over-eating, which decreased in many cases as the fat stores increased, only to reappear after starvation. However, he drew no conclusions as to the physiological mechanism involved. As he says (1946): 'Expressed as a problem in thermodynamics, the situation appears to be fairly clear—the more the animal eats, the more obese it becomes. But, expressed in physiological terms as a problem in the regulation of energy exchange, the situation is still obscure. The question, 'Why does the animal with hypothalamic lesions eat so much extra food?' immediately raises the questions, 'Why does the normal animal eat food? What determines how much it will eat? What changes in the internal environment set the animal to eating, and what changes are associated with satiety?' A partial answer to these questions seems to be offered by a study of the differences in reaction shown by thin and fat animals in our feeding experiments. First, is hyperphagia really the essential feature of the disturbance caused by hypothalamic lesions, as is suggested by Brobeck, and by the many descriptions of 'enhanced appetite', 'voracious, tigerish appetite', 'high degree of voracity', and so on, in the literature? Such behaviour does occur postoperatively, or after starvation, but in fully developed obesity the excess food intake is never striking. Certainly the animals show no sign of voracity. In our experiments the degree, and even the nature of the reaction to a change of diet, was conditioned by qualities in the food which were almost without effect in a young, normal animal. Far from being voraciously omnivorous, the fat animals appeared to be more than usually discriminative, the discrimination increasing with the obesity. Thus, the overall effect of the hypothalamic lesion seemed to be to emphasize two components of the urge to eat, a primitive voracious one, and a more discriminative component, either of which might be dominant according to the state of the fat stores. The possibility that the rejection of the kaolin diets by the fat animals was due to the gut being distended by the diluent is slight, as some refused the diets altogether for several days, and few ate enough at first to produce even normal distension. On the other hand, they were capable of an intake of

a more attractive diet amounting to two or three times normal. The rejection appeared to depend upon smell, and taste, and to be a conditioned response of the type usually described as appetite.

In marked contrast was the behaviour of the young rats. Adolph, who concentrated on these, concluded: 'Food acceptance and the urge to eat in rats are found to have relatively little to do with a local condition of the gastro-intestinal canal, little to do with the organs of taste, and very much to do with quantitative deficiencies of currently metabolized materials.' We confirmed that a chemical control of food acceptance did appear to be dominant in the 200 g. rat. Refusal of food on the presentation of unpalatable diets very quickly resulted in a degree of hunger sufficient to compel the acceptance of the diet. Excessive intake, which would have resulted in obesity, could not be induced. The emergence of the characteristic responses of the hyperphagic animal was seen in a proportion of the older unoperated rats studied, apparently the 'naturally' fat ones. It seems reasonable to assume that the superficially simpler feeding reactions of the young rat are the result of closer co-ordination of the different aspects of food acceptance rather than a real difference in mechanism.

The apparent differences in behaviour may be explained on the following hypothesis. The hyperphagia exhibited by rats with hypothalamic lesions appears to represent a primitive unconditioned urge to eat, which may be described as hunger. Normally this is inhibited by a hypothalamic satiety mechanism, which is directly sensitive to changes in the blood brought about by the ingestion of food. This hypothalamic mechanism appears to be particularly effective in the young rat, but becomes less well co-ordinated with increasing age, or after operative damage, and the failure in control of blood metabolites is then reflected in wider variability in fat depots. In such rats the palatability of food has an effect on the weight considerably greater than is seen in the young animal, and becomes the principal determinant of the degrees of obesity on unrestricted feeding. The feeding reactions of naturally fat animals are so similar to those after hypothalamic operation as to suggest that the difference is only one of degree, and that modification of the hypothalamic control of energy intake may be a prerequisite of obesity in general.

REFERENCES

Adolph, E. F. 1947 *Amer. J. Physiol.* 151, 110.
Babinski, M. J. 1900 *Rev. neurol.* 8, 531.
Brobeck, J. R. 1946 *Physiol. Rev.* 26, 541.
Brobeck, J. R., Tepperman, J. & Long, C. N. H. 1943 *Yale J. Biol. Med.* 15, 831.
Camus, J. & Roussy, G. 1913 *C. R. Soc. Biol., Paris,* 75, 483.
Camus, J. & Roussy, G. 1920 *Rev. neurol.* 36, 1201.

Camus, J. & Roussy, G. 1922 *Rev. neurol.* **38**, 622.

Erdheim, J. 1904 *S. B. Akad. Wiss. Wien*, Abt. III, **113**, 537.

Fröhlich, A. 1901 *Wien. klin. Rdsch.* **15**, 882, 906.

Hetherington, A. W. 1943 *Amer. J. Physiol.* **140**, 89.

Hetherington, A. W. & Ranson, S. W. 1939 *Proc. Soc. Exp. Biol., N.Y.*, **41**, 465.

Hetherington, A. W. & Ranson, S. W. 1942 *Amer. J. Physiol.* **136**, 609.

Mohr, —. 1840 *Wschr. ges. Heilk.* p. 565 (quoted by Brobeck).

Smith, P. E. 1927 *J. Amer. Med. Ass.* **88**, 158.

Smith, P. E. 1930 *Amer. J. Anat.* **45**, 205.

The adrenal homologues in the lungfish *Protopterus*

By W. Holmes

Department of Zoology and Comparative Anatomy, University of Oxford

[Plates 26 to 28]

No tissue representing the cortex of the adrenal gland has yet been described in the Dipnoi, though it is known in elasmobranchs and in all tetrapod vertebrates.

In the mammalian adrenal, lipine-containing inclusions give the cortical cells a characteristic appearance at certain stages of their life history. All those viscera of *Protopterus* which might be suspected of containing cortical tissue were studied in sections by a histochemical technique specific for phospholipines. Large intracellular droplets containing phospholipine were demonstrated in a tissue widely distributed around the kidneys, gonads and dorsal aorta throughout the body cavity.

The medullary homologue was identified by the chromaffin reaction, and proved to lie, as stated by Giacomini, in the walls of the intercostal branches of the dorsal aorta. The innervation of these medullary cells, from the sympathetic chains, was demonstrated by a silver method.

It is suggested that the lipine-containing tissue is that which became the cortex of tetrapods. Its distribution in *Protopterus*, and its relations with the medullary cells, are such that the elasmobranch and tetrapod adrenals could be derived from it by varying degrees of suppression and migration of the tissues.

Amongst Amphibia the adrenal of the Gymnophiona is most similar in arrangement to that of *Protopterus*.

The lipine tissue is so situated as to be readily available for biochemical and endocrinological studies.

INTRODUCTION

It has long been accepted that the two histologically distinct components of the adrenal gland of tetrapods, cortex and medulla, are represented in the elasmobranch fishes by tissues that are anatomically separate. The cortex is homologous with the interrenal organ of elasmobranchs, a structure which may be median, single and elongated (*Scyllium* type); paired, and broken up (*Raja* type); or asymmetrical, single and condensed (*Torpedo* type) (Dittus 1940). In all these cases the interrenal tissue is confined to the posterior region of the body cavity. The adrenal medulla of the tetrapods is homologous with the paraganglia of elasmobranchs; these are groups of cells sometimes large enough to be seen in a dissection, sometimes only

identifiable by histological investigation. They are found, in *Scyllium* for example, immediately dorsal to the posterior cardinal sinus, along the whole length of the body cavity. The anterior paraganglia are irregularly arranged, while those in the kidney region follow a regular segmental pattern, and are embedded in the dorsal surface of the kidney (Young 1933). Paraganglia are always found in a more or less close relationship with the ganglia of the sympathetic nervous system, as their name implies. The evidence that the interrenal and paraganglia are the homologues of cortex and medulla is reviewed by van der Sprenkel (1934).

Students of comparative anatomy must often have speculated as to how the elasmobranch condition of the adrenal could have been transformed in the course of evolution into that to be found in the tetrapods. But since it is now believed that the living elasmobranchs are much less primitive fishes than was formerly supposed, it can no longer be expected that their structure would represent directly the characteristics of the ancestors of the land-living vertebrates. And the condition of the adrenal in the actinopterygian fishes has been the subject of unsettled controversy ever since it was first investigated; the evidence brought forward from this group throws no light on the evolution of the tetrapod organ (van der Sprenkel 1934).

It is natural in these circumstances to turn to the crossopterygian fishes, to which group the tetrapod ancestor is believed to have belonged. And apart from the elusive *Latimeria* the only living representatives of the Crossopterygii are the Dipnoi, the lungfishes.

The dipnoan adrenal homologues were investigated in the African lungfish, *Protopterus*, by Giacomini (1906). His results are recorded in a paper described as a preliminary note. He concluded that there is no interrenal; that is to say, no cortical homologue, in the lungfish, and that the suprarenal component, the medullary tissue, is represented by nests of cells which lie around each intercostal artery in the trunk, and by similar nests in the wall of the more anterior parts of the left posterior cardinal vein, up to the ductus Cuvieri. It is unfortunate that Giacomini never published a complete account of his work, for as it stands, with no practical details, and without illustrations of any kind, it forms a most insecure foundation for discussion of the history of the adrenal.

MATERIAL

I am indebted to three zoologists who presented me with the specimens of *Protopterus* with which this work was done. Professor J. Gray of Cambridge generously gave me a fish in its cocoon, and on being hatched in Oxford it proved to be a lively specimen 20 cm. in overall length. Dr R. S. A. Beauchamp of the Jinja Laboratory of the East African Fisheries Research Organization kindly made special arrangements to send me six smaller living specimens by air from Uganda, and these with the larger animal lived apparently healthy lives in fresh-water tanks at 30° C. They were kept separate from each other, because of their ferocious habits, and fed well on small living molluscs, earthworms, and *Lumbriculus*. When killed as required each seemed in good health, and one was allowed to survive for eight

months. Preserved material was obtained for me by Mr J. R. Clarke, on the Oxford University Expedition to the Gambia, 1948, and I am grateful for the care with which he fixed the material according to my instructions. These animals were some 25 cm. long, and were more useful for some purposes than the smaller specimens.

Of the three species of *Protopterus* distinguished by Holly (1933) mine would appear to be *P. annectens*. However, none of my specimens were fully grown, and since the distinction between *P. annectens* and *P. aethiopicus* turns on relative external proportions, which varied according to the size of the animal, I am not able to be certain on this point.

Some of the fish were killed by decapitation, others by chloroform vapour after preliminary anaesthesia in a urethane solution. The body cavity was opened ventrally by a median incision from the level of the heart to a point posterior to the cloaca, cutting through the pectoral and pelvic girdles. The gut and liver were then removed. The appearance of a dissection of a female at this stage is shown in figure 1. The kidneys are asymmetrical anteriorly, for that on the morphological right side ends in intimate contact with the liver. The posterior vena cava, which passes more or less through the substance of the liver, is shown in the drawing only in its course along the right kidney, for it is cut, and its anterior part removed, when the liver is cut out. The pulmonary vein and the ventral branch of the pulmonary artery are seen passing along the ventral surface of the paired lungs, which extend some three-quarters of the distance from pectoral to pelvic girdles. The dorsal branch of the pulmonary artery is not visible. The left posterior cardinal vein is not shown where it passes from the tip of the left kidney to the ductus Cuvieri. Here it is a much finer vessel than on its course along the kidney, and may well have an insignificant part in the venous return to the heart, for there are at least two large anastomotic vessels from the left kidney across to the inferior vena cava, which is on the right side. The dorsal aorta is, of course, single and median, lying dorsal to, and between, the lungs. It is intimately bound to the ventral sheath of the notochord.

The anatomy of *Protopterus* has already been thoroughly explored, and since no adrenal cortical tissue has been discovered it was clear that no further advance could be obtained by dissections alone.

<div align="center">METHODS AND RESULTS</div>

The direct demonstration that a given structure is an organ of internal secretion requires that it should have been located anatomically and that it is possible to remove it experimentally. Or it may be found in a pathological state co-existent with abnormal physiological phenomena in the animal. The whole repertoire of experimental methods of endocrinology can then be brought into play. But these direct methods cannot be applied to an animal in which the suspected cells or tissues are unknown, where there is no definite organ to which the role in question could be attributed. *Protopterus* is such a case, and the adrenal homologues must be sought by other means. Nor is the embryological criterion of homology applicable, since it requires that a structure recognized in its fully developed form should be

traced back to its embryonic rudiments. In the case of the adrenal cortical element at least, there is no evidence of its existence in the adult that would allow of this approach. We turn, then, to the establishment of a cytological or histochemical homology, attempting to find in the adult lungfish cells or tissues which have characters sufficiently in common with known tetrapod adrenal cells, and with no other sort of cell, as the first stage in the investigation.
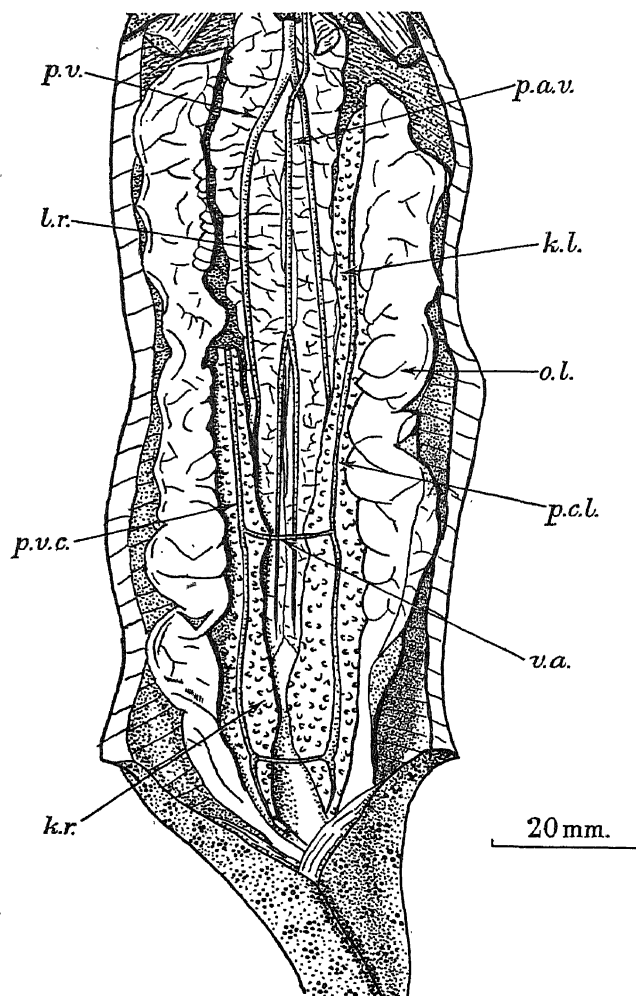
FIGURE 1. *Protopterus*. Female dissected from the ventral side. Post-cardinal viscera as seen after the removal of the alimentary canal and liver. *k.l.* left kidney; *k.r.* right kidney; *l.r.* right lung; *o.l.* left ovary; *p.a.v.* ventral branch of pulmonary artery; *p.c.l.* left posterior cardinal vein; *p.v.* pulmonary vein; *p.v.c.* posterior vena cava; *v.a.* venous cross-anastomosis.

The cortical tissue

The cells of the mammalian adrenal cortex cannot be characterized by a single cytological description, but their lipoid inclusions give them an appearance 'possessed normally by no other cells in the body, with the exception of certain

stages of the corpus luteum' (Hoerr 1936*b*). They present a variety of appearances differing not only between the principal zones into which the cortex is divided, but also within these zones themselves. It is now clear that this variety is a manifestation of the occurrence of phases in the process of cell growth and secretion; it can be modified in a variety of experimental situations (e.g. Zwemer, Wotton & Norkus 1938; Bennett 1940). Histochemical studies of the normal gland show that intracellular droplets containing phospholipine appear, grow and disappear, and so do other lipoid droplets, negative to phospholipine tests. (The terms 'lipoid' and 'lipine' are used here with the meanings defined by Cain (1950).) The relation of these droplets to the adrenal cortical hormones is still not clear.

Students of the adrenal cortex of non-mammalian vertebrates have found in those animals also similar cell inclusions, though most of the investigations were made before lipoid histochemistry had reached its present-day precision, as was, for example, the work of Fraser (1929) on the elasmobranch interrenal. These reports, summarized by van der Sprenkel (1934), were usually based either on preparations revealing the anisotropy of the inclusions, or on material fixed in solutions containing osmium tetroxide. It was by the latter method that Stenger & Charipper (1946) followed the ontogeny of the frog cortical tissue. But the unreliability of the osmium technique as a histochemical method is now well known (Hoerr 1936*a*).

The lipine test

Droplets containing phospholipine of the size and number found in the mammalian cortical cells at certain stages of their life history give these cells a unique appearance. The plan of the present search for the cortical homologue in the lungfish was to apply Baker's acid haematein test for phospholipine (1946) to all those organs and tissues in the body cavity posterior to the heart in which it seemed possible that the homologue, if existing, might be located. The adrenal tissue of the frog was treated by an identical procedure, so as to enable direct comparisons to be made. This is, of course, only a tentative preliminary step in the investigation. A tissue as complicated as we believe the adrenal cortex to be, in mammals at least, cannot be studied adequately in a few animals of unknown age and in a species whose normal metabolism is so little known.

In a lungfish dissected as far as in figure 1, the remaining viscera, including the venous vessels, pulmonary vagi, dorsal aorta and sympathetic chains, are held together in a common investing connective tissue membrane. If, therefore, the gonad, lung and kidney of each side are gently raised, and freed from their delicate connexion to the coelomic wall at their outer edge, the group of organs is held in place only by the tissue which binds the dorsal aorta to the fibrous sheath of the notochord in the middle line. The next step presents some difficulty, since the intercostal branches of the aorta pass up and around the notochord into the dorsal musculature (figure 12, plate 27), and their preservation is important in view of Giacomini's conclusions. Furthermore, the sympathetic nerve chains lie close on either side of the aorta, and loop round the intercostal branches (figure 2; figure 21, plate 28). But with care the specimen of viscera, with aorta and the origins of the intercostals and sympathetic, can be removed entire.

Baker's acid haematein test was applied exactly according to his directions (1946). It requires first of all fixation for six hours in formol-calcium. In the case of one fish this was done with the viscera *in situ*, to avoid trauma to the fresh tissues; in two other animals the viscera was removed before fixation, with the compensating advantage of a more ready penetration of the fixative.

Since Baker's test is applicable only to frozen sections, and since the texture of the tissues made it necessary for them to be embedded in gelatin before sectioning, there was no possibility of studying serial sections. The plan was instead to make first a general survey of the material. Transverse and longitudinal sections were cut from all the blocks of material from one animal, and it was intended to use the other two individuals for a detailed study of any region found to be of interest. In fact, however, the second and third batches of material served only to give confirmation of what was found in the first, for large droplets containing phospholipine were found in the cells of a tissue that is widely distributed amongst the viscera. The sections were differentiated in borax-ferricyanide for the maximum period specified by Baker: 18 hours. The cells whose inclusions reacted positively to Baker's test will henceforward be referred to as lipine cells. They are very large, as are so many cell types in the Dipnoi, as compared with similar cells in other vertebrates. Figure 4, plate 26, shows the lipine cells (cortical cells) of the frog adrenal at the same scale as figure 5, plate 26, which shows lungfish lipine cells. Figures 9 and 8, plate 26, are of the same cells, frog and lungfish respectively, as seen with an oil-immersion objective. And these figures also show that the lipine droplets are much larger in the fish, although their arrangement and number is similar to that in the amphibian. In any field of an acid haematein preparation of the fish material in which lipine cells are present they lie adjacent to others in which droplets can be seen as a 'negative' picture, being represented only by what seem to be vacuoles of similar size to the droplets, but unstained against a pale background of cytoplasm (figures 5 and 8, plate 26). Since the lipine cells are so large, often more than $20\,\mu$ in diameter, it is not impossible that the unstained vacuoles are in fact droplets from which the stain has been removed by overdifferentiation. For $15\,\mu$ sections will expose the cell inclusions to the action of the differentiating fluid to varying degrees. On the other hand it is perhaps more likely that the cells negative to the test do in fact lack phospholipine inclusions, and are cells at a different stage in the cycle of secretion. Their similarity to the lipine cells is such that it would be unjustifiable to distinguish them as a different type of cell.

In a minor proportion of the lipine cells the inclusions are smaller than usual, leaving much more unstained cytoplasm between them (figure 7, plate 26). This may be taken as further evidence of a secretory cycle in the tissue. Sometimes the lipine droplets appear to be extracellular, released into the blood vessels, for example, but this may well be a traumatic artefact.

The lipine cells, intermingled with those that are vacuolated but negative to the acid haematein test, are widely distributed amongst the viscera examined. It was possible to survey their distribution in some detail, for once identified in the specifically stained material they proved to be recognizable in serial sections fixed for other purposes. For in the search for the adrenal medullary homologues, and

for the sympathetic-medullary cell relationships, I made use of serial paraffin sections of material fixed in formol-potassium dichromate, and formol-mercuric chloride. In material of this kind it was found that the tissue containing lipine cells was still distinguishable, for the fixatives left vacuoles recognizable even when only a nuclear stain had been employed on the sections (figure 23, plate 28). And in silver preparations they were often made very clear by a 'metachromatic' staining of their walls or of the residuum of their contents (figures 20 and 21, plate 28).
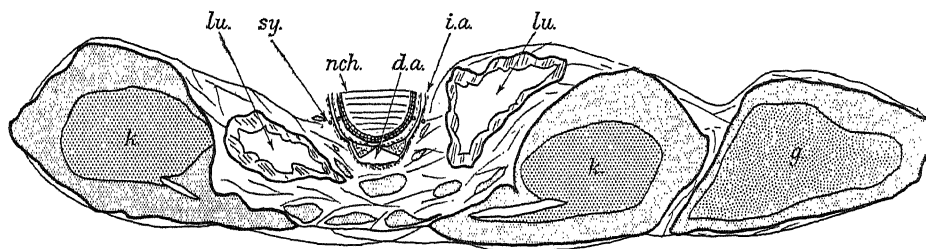
FIGURE 2. *Protopterus*. Diagrammatic transverse section in abdominal region, showing distribution of lipine cells in relation to kidney, lungs, gonad and dorsal aorta. Left gonad not shown. Lipine tissue finely stippled. *d.a.* dorsal aorta; *g.* gonad; *i.a.* intercostal artery; *k.* kidney; *lu.* lung cavity; *nch.* notochord; *sy.* sympathetic cord.

Identified in this way the lipine cells lie principally in the epithelial coelomic connective tissue which suspends the viscera investigated. Isolated, or in small groups, they are scattered throughout this tissue, and are thus often found very close to the aorta and sympathetic chain (figure 20, plate 28). At various points they are found in masses, particularly on the dorsal surface of the kidney, and they form a substantial part of what is, anatomically, the kidney (figure 2; figure 5, plate 26). Here they form the 'lymphoid tissue' of Parker (1892). This covering to the kidney, present but thinner on its ventral surface, extends also over the dorsal surface of the gonad, but may not surround it completely. Aggregates are also found between the kidneys and, particularly in the posterior region of the body, solid masses of lipine cells lie in a position between and somewhat dorsal to the kidneys, in a position corresponding with that of the median selachian interrenal (figure 2; figure 11, plate 27); here they are sometimes associated with adipose tissue. These median 'interrenal' masses were observed by Parker, who noted that they had the same cytological characteristics as the 'lymphoid' tissue round the kidney (Parker 1892, p. 184). It seems that the small aggregates of lipine cells close to the intercostal arteries and sympathetic are comparable with the groups of cells noted by Jenkin (1928) in *Lepidosiren* and again described as lymphoid tissue (Jenkin's figure 4).

Lipine cells are also to be found scattered sparsely within the kidney tissue itself amongst the tubules, as are the cortical cells of the frog (figure 4, plate 26) (Beatty 1940), and some lie on the outer walls of the lungs.

The 'cholesterol' test

The validity of Liebermann's test for cholesterol, and its various modifications, has often been discussed (see Cain 1950). For our present purpose it suffices to

point out that a positive reaction is given by the adrenal cortical cells of mammals (Yoffey & Baxter 1949), since cholesterol, or its esters, or the related substances which give a positive reaction to the test, are a distinguishing characteristic of these cells. The following modification of Liebermann's test, 'controlled' on mammalian kidney, was applied to the lungfish lipine cells. A portion of kidney with its lipine cells was fixed in saline formaldehyde, and frozen sections were cut at 15μ. The sections were treated with a 2·5 % solution of iron alum at 37° C for 24 hours. They were then rinsed in distilled water, floated in a few drops of water on a slide, and the surplus water removed with filter paper so as to leave the section flat on the slide. Two or three drops of a mixture of equal parts of pure sulphuric and acetic acids were then placed on the section, and a coverslip lowered on to the liquid. A positive reaction to this test is shown by a dark green coloration.

In every section to which the test was applied numerous groups of lipine cells gave an unmistakeably positive reaction (figure 10, plate 27). But the majority of the cells gave a negative result. There was no recognizable special location of the positively reacting cells; they were found in all positions, though always in groups of at least three. Nor did they appear to be distinguished from the other lipine cells in any way other than by their reaction to the cholesterol test, so far as could be judged from the preparations, which are such as to be unsuited to detailed cytological examination.

The chromaffin tissue

Cells which produce adrenalin are found in mammals in the adrenal medulla, and in scattered groups called 'chromaffin glands' or paraganglia, variously distributed in the visceral and coelomic walls. The cytological identification of adrenalin cells, which preceded the demonstration of their hormone content, depends on the well-known chromaffin reaction. The discovery that the cytoplasm of adrenal medullary cells shows a brown or yellow coloration in material fixed in solutions containing potassium dichromate is of considerable antiquity. Many workers satisfied themselves by controlled experiments that the chromaffin reaction is a valid histochemical test for adrenalin (e.g. Kingsbury 1911; Ogata & Ogata 1923). It remained, however, for Gerard, Cordier & Lison (1930) to prove that the reaction is not specific to adrenalin, but is a general one for polyphenols, aminophenols, and certain polyamines. The reaction is equally well given by fixation in fluids containing, for example, potassium iodate instead of dichromate; it depends on the formation of a coloured compound by the oxidation of the substrate. But in spite of this qualification of the specificity of the test it remains the most useful and most specific method for the detection of adrenalin cells. The discovery of the accessory adrenalin-producing paraganglia in the mammalian coelomic wall would have been impossible without it.

Since trauma is known to affect the distribution of adrenalin in tissues, the same procedure was adopted here as with the lipine cell investigation; that is to say, material from some animals was fixed *in situ*, from others after dissection. Overanaesthesia and decapitation were used as alternative methods of killing the animals. I was unable to find any difference between the material obtained in these various ways.

The tissues were fixed for 24 hours in:

40 % formaldehyde	10 ml.
3 % potassium dichromate	90 ml.
sodium chloride	0·7 g.

Some material was then transferred to a 3 % potassium dichromate solution for a further 2 days, washed in running water, embedded and sectioned. Other specimens were taken from the formol-dichromate to plain formol saline for 2 days before washing and embedding. These variations also seemed to make no difference to the appearance and distribution of recognizable chromaffin cells.

The total length of the excised visceral specimen was of the order of 9 cm. in each case. Large numbers of serial sections were made and studied, so as to give a reasonably complete picture of the specimen as a whole. The nuclei in the sections were stained with Mayer's haemalum.

Interpretation of the sections

Inspection of any of the sections made as described above reveals at once the presence of an abundance of cells containing yellow, brown and black granules variously, particularly in the kidney and coelomic epithelia. But control sections, of material fixed in formol-mercuric chloride and stained with haemalum alone, show that this pigmentation is equally evident in the corresponding cells, though the tissue has not had dichromate treatment (figure 23, plate 28). It is clear that the lungfish tissues are abundantly pigmented in nature, and caution is needed in interpreting brown granulation as evidence of adrenalin content.

Many accounts of the cytology of chromaffin cells that are to be found in the literature make it clear that their authors found the chromaffin reaction confined to cytoplasmic granules in the tissues they were studying; some, indeed, give drawings (though never photographs) of very coarse brown granules found in dichromate-fixed material (e.g. Sacarrão 1944, on the suprarenal of the dogfish). I was led to doubt this account of the cytology of these cells first because I found in chromaffin preparations of the frog adrenal that the cytoplasm of the medullary cells was a homogeneous yellow-brown in colour, with no granulation comparable in size with that often figured. And Lison (1931), whose full discussion of the cytology of the chromaffin reaction inspires great confidence, states clearly that it is manifested as a uniform, non-granular cytoplasmic coloration. And since all cells in my lungfish preparations in which brown granules were to be found were equally brown in preparations untreated by dichromate I was led to believe that I should expect a true chromaffin reaction to manifest itself as a cytoplasmic rather than granular stain. It is to be borne in mind that many students of the reaction employed chrome fixatives which contained also protein precipitants.

A further pitfall in interpretation lies in the fact that the cytoplasm of the red blood cells of the lungfish shows a marked, cytoplasmic, chromaffin reaction, and every red cell has a yellow-brown cytoplasm in the dichromate preparations (figure 14, plate 27). It is easy to distinguish a red cell when it lies alone, or with a group of similar cells, in the lumen of a vessel. But it is often the case that the

corpuscles lie intimately amongst the endothelial cells of the vessels, and they are, of course, in the lungs, constantly found in intimate association with the membranous walls of the pulmonary spaces. Similarly they are regularly found in the renal capillaries, so distorted in the small vessels they occupy, that they seem to be an element of the kidney's structure. However, the use of Mayer's haemalum as a counterstain, and care in interpretation, enable red cells to be distinguished with some certainty, and excluded in the search for the adrenal medullary cells. For the nuclear structure and staining reaction, the latter being more feeble than that of other cells in carefully stained sections, are diagnostic of the red cell in all the ambiguous situations which it occupies.

Yet another possible source of confusion arises from the fact that in the chromed material the cytoplasmic droplets of the lipine cells have a yellowish brown colour; indeed, they are faintly yellowish in all material, as might be expected from their size and nature. But a previous knowledge of the distribution of such cells enables them to be excluded from the search.

When all these possible sources of error have been taken into account there remains only one type of cell which seems to fulfil all the requirements for identification as the true chromaffin adrenal-containing cell. This is located in the walls of the intercostal branches of the dorsal aorta. These chromaffin cells extend in this position from a point shortly after the branch has left the aorta along some two-thirds of its length before it bifurcates into dorsal and ventral branches; that is, between the points marked A and B in figure 12, plate 27.

The intercostal arteries arise at intervals from both sides of the dorsal aorta, one corresponding approximately to each myotome. The branches do not, however, come off with perfect regularity, and that on one side of the aorta often arises at a slightly different level from that on the other side.

The cytoplasm of the chromaffin cells is distinguished by its coloration, yellow with a tinge of brown in sections at $8\,\mu$. The whole of the cytoplasm is uniformly coloured and appears homogeneous, without granulation. In some cases it forms a uniform ring round the nucleus; in others the nucleus is asymmetrically placed, and the cytoplasm of the side of the cell farthest from the nucleus is prolonged into strands which have a characteristic 'dendritic' appearance, and often point towards the lumen of the artery (figures 13 and 14, plate 27); the chromaffin cells of the frog adrenal have a closely similar appearance. The lungfish chromaffin tissue lies in the wall of the artery, often immediately adjacent to the endothelium (figures 13, 14 and 15, plate 27); in other cases, perhaps only in the larger animals, it lies as a small cellular nest separated from the intima of the vessel by one or two layers of the cells of the vessel wall (figures 16 to 19, plate 28). In the hope of making possible a more precise statement about the location of the tissue in relation to the coats of the artery, some of the sections were counterstained by Verhoeff's technique for elastic fibrils. It was found, however, that though elastic fibres are abundant in the two blocks of tissue lying on the dorsal aspect of the aorta (figure 11, plate 27), there was no elastic lamina in the intercostal artery (in an animal 12 cm. in length). In the smallest lungfish the intercostal arterial wall seemed to be little more than one cell thick, and the chromaffin cells were applied against this as a flat sheet, also

one cell thick (figure 11). I could find no chromaffin tissue in the walls of any of the venous trunks posterior to the heart.

The innervation of the chromaffin tissue

Numerous investigations of the mammalian adrenal have made clear that the medullary cells have a profuse sympathetic innervation, while the cortical cells have no direct nerve supply from any source (e.g. McFarland & Davenport 1941; McFarland 1944). A comparable state of affairs is found in the elasmobranchs, the paraganglia being supplied from the adjacent sympathetic tissue, the interrenal uninnervated (Young 1933).

The chromaffin-sympathetic relations were examined in the lungfish in material fixed in:

40 % formaldehyde	10 ml.
mercuric chloride, sat. aq. sol.	90 ml.

The nerve axons were demonstrated by the buffered silver-pyridine technique (Holmes 1947).

The sympathetic nerve chain of *Protopterus* is very small in diameter, and in my largest Gambia specimen measured only 40μ; in an animal 15 cm. long the chain was 25μ in diameter (figure 20, plate 28). It lies in the sheath which binds the lungs to the dorsal body wall, and is of course double, one sympathetic lying on each side of the aorta. The sympathetic forms a loop round each intercostal artery, and at this level has no definite outer sheath (figure 2; figure 21, plate 28); the sympathetic-intercostal relations are just as was described in *Lepidosiren* (Jenkin 1928). There are no anatomical ganglia in the chain in the regions examined; ganglion cells lie in the cord at all levels; often they are collected into small groups (figure 16, plate 28), but these are never so large as to form a swelling on the nerve.

The silver-stained preparations show that the chromaffin tissue in the arterial wall always receives a nerve supply from the sympathetic cord which surrounds it. The axons innervating the tissue often come off singly (figure 16). There seems to be no anatomical regularity in their course amongst the chromaffin cells, but there is no doubt that an intimate relationship is established; one axon serves several cells *en passant* (figure 19, plate 28). I have not been able to distinguish any definite characteristic features of the termination of the axon on the surface of the cell it supplies.

The sections were cut at 15 or 20 μ, and in this material axons can be followed for considerable distances by continuous up-and-down focusing of the microscope. But photographs are confined to one focal plane, and those reproduced in plate 28 give a more fragmentary picture of the chromaffin innervation than is obtainable by direct inspection of the material. Figures 16 to 19 are of the chromaffins of one particular intercostal vessel, photographed in adjacent sections. The series came from one of the Gambia specimens, whose larger size made the innervation more clearly demonstrable. Unfortunately as a result of the inevitably prolonged fixation of the material the cytoplasm of the chromaffins is heavily stained with silver, and there are some non-nervous fibrous structures in the sections. Nevertheless the

axons can be distinguished. Sections of a smaller animal, fixed for three days only, are more satisfactorily stained (figures 21 and 22), but the extremely fine diameter of the sympathetic fibres makes them less suitable for the study of the relationships at issue. I have not seen any significant number of nerve axons in the walls of blood vessels other than the intercostals. It is not possible to draw any conclusions as to whether the chromaffin innervation is pre- or post-ganglionic.

DISCUSSION

It is suggested here as a hypothesis that the lipine-containing tissue is *Protopterus* is the adrenal cortical homologue. Evidence for and against this hypothesis can be gained not only by further studies of the lungfish, but also by the comparative histochemical study of the gland in other vertebrates. It is evident that some new approach such as this is necessary to make clear, for example, the condition of the gland in Teleostei. That the necessary information is not already available is due, in part, to the difficulties of lipoid histochemistry: the acid haematein test, for example, is the only fully reliable method for distinguishing phospholipines in sections, and is of recent discovery (Baker 1946, 1947).

This paper of course contains no evidence that the lipine tissue has an endocrine function, but its distribution and relation to the chromaffin cells is particularly suggestive. For if the homologies are accepted, then by suppression of cortical cells in varying areas, and migration of medullary cells, the lungfish condition could have led to that found in elasmobranchs, Amphibia, and higher tetrapods (figure 3).

The interrenal tissue in metamorphosed anurans is almost entirely confined to the kidney region, and the majority of the chromaffin cells are similarly restricted in distribution, being intermingled with the cortical cells to form the coloured adrenal bodies on the ventral surface of the kidney (figure 3). In urodeles the interrenal tissue extends throughout the abdomen, usually in separate masses, though these in places are fused together. The chromaffin cells are similarly distributed, being adjacent to the interrenal cells in the cranial region, intermixed with them in the kidney region. In the Gymnophiona the cranial part of the cortical tissue is fused to a single mass, but caudally it is broken up into patches (figure 3; Dittus 1936). In the elasmobranchs the cortical tissue is confined to the kidney region, showing various degrees of concentration and fusion, most extreme in the *Scyllium* type (figure 3). It is clear that the condition in all Amphibia and that in the elasmobranchs could be divergently derived from the more generalized type as repesented in the lungfish *Protopterus*. It should perhaps be emphasized that the abundance of the lipine tissue, and the ease with which it could be obtained uncontaminated in considerable quantities, should facilitate physiological and biochemical investigation.

The finding of chromaffin cells in the walls of the intercostal arteries confirms Giacomini's discovery; the numerous sources of error in the interpretation of chromed sections makes it possible that he was mistaken in describing them in the walls of the left posterior cardinal.

That the intercostal chromaffins are the medullary homologue is supported by the demonstration of their sympathetic innervation.
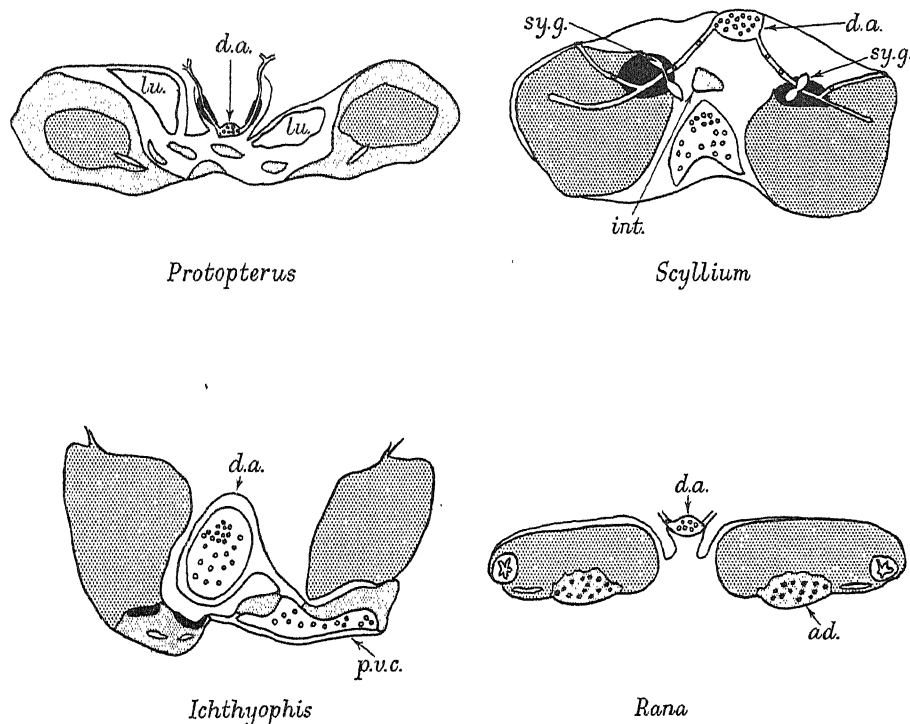
FIGURE 3. Adrenal cortical and medullary tissue in diagrammatic transverse section of the kidney region in: *Protopterus* (Dipnoi); *Scyllium* (Elasmobranchii); *Ichthyophis* (Amphibia, Gymnophiona); *Rana* (Amphibia, Anura). Cortical tissue—fine stippling; medullary tissue—black; kidney—coarse stippling. *ad.* adrenal; *d.a.* dorsal aorta; *int.* interrenal; *lu.* lung cavity; *p.v.c.* posterior vena cava; *sy.g.* sympathetic ganglion. (*Scyllium* after Young 1933; *Ichthyophis* after Dittus 1936.)

REFERENCES

Baker, J. R. 1946 *Quart. J. Micr. Sci.* **87,** 441.

Baker, J. R. 1947 *Quart. J. Micr. Sci.* **88,** 463.

Beatty, B. 1940 *Proc. Leeds Phil. Lit. Soc.,* Sci. Sect., **3,** 633.

Bennett, H. S. 1940 *Amer. J. Anat.* **67,** 151.

Cain, A. J. 1950 *Biol. Rev.* **25,** 73.

Dittus, P. 1936 *Z. wiss. Zool.* **147,** 459.

Dittus, P. 1940 *Z. wiss. Zool.* **154,** 40.

Fraser, A. H. H. 1929 *Quart. J. Micr. Sci.* **73,** 121.

Gerard, P., Cordier, R. & Lison, L. 1930 *Bull. Histol. Tech. Micr.* **7,** 133.

Giacomini, E. 1906 *R.C. Accad. Lincei,* **15,** 394.

Hoerr, N. L. 1936*a* *Anat. Rec.* **66,** 149.

Hoerr, N. L. 1936*b* *Anat. Rec.* **66,** 317.

Holly, M. 1933 *Das Tierreich,* Lief. 61. Leipzig: Akad. Verlag.

Holmes, W. 1947 'The peripheral nerve biopsy.' In *Recent advances in clinical pathology,* ed. S. C. Dyke, p. 402. London: Churchill.

Jenkin, P. M. 1928 *Proc. Roy. Soc. Edinb.* **48,** 55.

Kingsbury, B. F. 1911 *Anat. Rec.* **5,** 11.

Lison, L. 1931 *Arch. Biol., Paris,* **41,** 344.

McFarland, W. E. 1944 *J. Exp. Zool.* **95,** 345.

McFarland, W. E. & Davenport, H. A. 1941 *J. Comp. Neurol.* **75,** 219.

Ogata, T. & Ogata, A. 1923 *Beitr. path. Anat.* **71,** 376.

Parker, W. N. 1892 *Trans. R. Irish Acad.* **30,** 109.

Sacarrão, G. F. 1944 *Arch. portug. Sci. biol.* **7**, 135.

van der Sprenkel, H. B. 1934 *Handbuch der vergleichende Anatomie*, ed. Bolk *et al.* 2. Band, 1. Hälfte, p. 777. Berlin: Urban und Schwarzenberg.

Stenger, A. H. & Charipper, H. A. 1946 *J. Morphol.* **78**, 27.

Yoffey, J. M. & Baxter, J. S. 1949 *J. Anat., Lond.*, **83**, 89.

Young, J. Z. 1933 *Quart. J. Micr. Sci.* **75**, 571.

Zwemer, R. L., Wotton, R. M. & Norkus, M. G. 1938 *Anat. Rec.* **72**, 249.

DESCRIPTION OF PLATES 26 TO 28

All the photographs, except figures 4 and 9, are of *Protopterus* material. Lettering: *ax.* sympathetic axon amongst chromaffin cells; *b.c.* blood cells; *cr.s.* group of chromaffin cells; *d.a.* dorsal aorta; *i.a.* intercostal artery; *i.a.br.* point of branching of intercostal artery; *li.s.* group of lipine cells; *lu.* lung cavity; *nch.* notochord; *sy.* sympathetic chain.

PLATE 26

FIGURE 4. Frog. Portion of a section through the ventral surface of the kidney. Acid haematein test. Lipine cells of the adrenal are seen on the surface of the kidney and amongst the kidney tubules. For comparison with figure 5.

FIGURE 5. Portion of a section through the ventral surface of the kidney. Acid haematein test. Kidney tissue below, lipine tissue above.

FIGURE 6. Lipine cells from the coelomic connective tissue. Acid haematein test.

FIGURE 7. Lipine cells from the surface of the kidney. Acid haematein test.

FIGURE 8. Lipine cells from the surface of the kidney. Oil-immersion objective. Acid haematein test. Positively reaction cells, with stained lipine inclusions, lie amongst negatively reacting cells with apparently vacuolated cytoplasm.

FIGURE 9. Frog. Lipine cells from the adrenal. Acid haematein test. Same magnification as figure 8.

PLATE 27

FIGURE 10. Lipine cells from the surface of the kidney. Cholesterol test. Positively reacting cells, black, amongst negatively reacting ones.

FIGURE 11. Transverse section through the dorsal aorta, showing the origin of an intercostal artery on one side, chromaffin tissue, and 'interrenal' mass below. Chromaffin reaction.

FIGURE 12. Transverse section showing the relations of the dorsal aorta to the notochord and lung, and the origin and dorsal branching of an intercostal artery. Silver.

FIGURE 13. Transverse section of an intercostal artery. Chromaffin reaction.

FIGURE 14. Longitudinal section of an intercostal artery, tangential to the lumen, in which blood corpuscles are visible on the left. Chromaffin reaction.

FIGURE 15. Longitudinal section of an intercostal artery showing lumen, endothelium and chromaffin cells in its walls. Chromaffin reaction.

PLATE 28

FIGURES 16 to 19 are transverse sections taken at adjacent levels of one intercostal artery, showing the relations of the sympathetic axons to the chromaffin cells. Silver.

FIGURE 20. Transverse section of the sympathetic trunk showing adjacent lipine cells. Silver.

FIGURE 21. Longitudinal section showing bifurcation of the sympathetic trunk as it approaches the intercostal artery. Silver.

FIGURE 22. Axons amongst chromaffin cells in the wall of an intercostal artery. Lipine cell to the right. Silver.

FIGURE 23. Pigment masses amongst lipine cells on the surface of the kidney. Haemalum, eosin.

10   50μ

11   *li.s.*   200μ

12   *d.a.*   200μ
cr.s.

13   *cr.s.*   25μ

14   *b.c.*   25μ

15   *cr.s.*   25μ

INDEX TO VOLUME 137 (B)